

Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»



**Информационные технологии и системы 2018
(ИТС 2018)**

Материалы международной научной конференции
(БГУИР, Минск, Беларусь, 25 октября 2018)

**Information Technologies and Systems 2018
(ITS 2018)**

Proceeding of The International Conference
(BSUIR, Minsk, Belarus, 25th October 2018)

Минск БГУИР 2018

УДК 004
ББК 32.81
И74

Редакционная коллегия:

Л. Ю. Шилин (главный редактор), Д.П. Кукин, А. А. Иванюк, А. В. Марков, В. В. Голенков, А. А. Навроцкий, А. Ф. Трофимович, Н.С. Жилач, С. В. Колосов, В.С. Муха, А. Б. Гуринович (ответственный секретарь)

Информационные технологии и системы 2018 (ИТС 2018) =
И74 Information Tehnologies and Systems 2018 (ITS 2018) : материалы международной научной конференции, Минск, Беларусь, 25 октября 2018 г. / Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2018.
ISBN

Сборник включает прошедшие рецензирование доклады международной научной конференции «Информационные технологии и системы 2018» (ИТС 2018).

Сборник предназначен для преподавателей высших учебных заведений, научных сотрудников, студентов, аспирантов, магистрантов, а также для специалистов предприятий в сфере IT-технологий.

Материалы сборника одобрены организационным комитетом и печатаются в авторской редакции.

УДК 004
ББК 32.81

ISBN

«Белорусский государственный университет информатики и радиоэлектроники», 2018

ПРЕДИСЛОВИЕ

25 октября 2018 года в городе Минске прошла международная научная конференция «Информационные технологии и системы» (ИТС 2018), организованная факультетом информационных технологий и управления учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», базовой организации государств-участников СНГ по высшему образованию в области информатики и радиоэлектроники.

В течение дня на пленарном и пяти секционных заседаниях было представлено 137 докладов по следующим направлениям:

Секция «Системы управления»:

- моделирование систем управления;
- автоматизированные исследования систем управления;
- проектирование систем управления;
- микропроцессорные системы управления;
- информационные технологии в системах управления;
- системы передачи и защиты дискретной информации;
- системы телемеханики;
- элементы и устройства систем управления.

Секция «Интеллектуальные информационные технологии»:

- проектирование баз знаний;
- интеллектуальный поиск информации;
- интеллектуальные решатели задач;
- пользовательские интерфейсы интеллектуальных систем.

Секция «Проектирование встраиваемых систем»:

- проектирование и быстрое прототипирование встраиваемых систем на ПЛИС;
- методологии проектирования IP-компонент встраиваемых систем;
- программная и аппаратная защита IP-компонент;
- контролепригодное проектирование и тестирование встраиваемых систем;
- реконфигурируемые встраиваемые системы.

Секция «Системы обработки информации»:

- математические модели и методы обработки информации;
- статистические, численные, логические, графовые, нейросетевые и другие;
- анализ многомерных данных;
- методологические проблемы и технологии подготовки специалистов в области систем обработки информации;
- технологии и стандарты программирования и проектирования систем обработки информации;
- системы и средства совместной работы пользователей;
- защита информации в автоматизированных системах обработки информации.

Секция «Математическое моделирование и компьютерная графика»:

- методы решения уравнений математической физики;
- методы нелинейной оптимизации;
- модели и методы радиофизики и приборов СВЧ;
- модели и методы тепло и массопереноса;
- математические модели и методы компьютерной графики;
- аппаратное обеспечение компьютерной графики;
- алгоритмы и методы интерактивных программных продуктов.

В работе конференции приняли участие учёные и специалисты Республики Беларусь, Российской Федерации, Украины.

Докладчики представляли как высшие учебные заведения, в их числе: БГУ, БГУИР, БНТУ, БГТУ, БГЭУ, БрГТУ, ПГУ, МИУ, Военная академия РБ, Гомельский государственный технический университета им. П. О. Сухого, Гродненский государственный университет им. Янки Купалы, ГГУ им. Ф. Скорины, ГУВПО «Белорусско-Российский Университет», Могилевский государственный университет продовольствия, Luxembourg Institute of Health, Черкасский государственный технологический университет, НИ Томский государственный университет, СПбГЭТУ «ЛЭТИ», Винницкий финансово-экономический университет – так и институты и коммерческие организации: НАН РБ, Объединенный институт проблем информатики НАН РБ, ИДСТУ СО РАН, ОАО «АГАТ-системы управления», филиал АО «СО ЕЭС» «Региональное диспетчерское управление энергосистемы Иркутской области», EPAM Systems.

Оргкомитет благодарит всех, кто счел возможным поделиться результатами своей деятельности и принял участие в работе конференции.

ORGANISING COMMITTEE

Prof.Dr.	Vadim Bogush	Rector of The Belarusian State University of Informatics and Radioelectronics
Dr.	Sergey Dick	First Vice-Rector of BSUIR
Dr.	Anatoly Osipov	Vice-Rector for Research of BSUIR
Dr.	Boris Nikulshin	Vice-Rector for Education of BSUIR
Dr.	Helena Zhivitskaya	Vice-Rector for Education of BSUIR
Prof.Dr.	Leonid Shilin	Dean of The Faculty of Information Technologies and Control of BSUIR
	Dmitry Kuznetsov	Vice-rector for Personal Development of BSUIR

PROGRAMME COMMITTEE

Prof.Dr.	Vladimir Golenkov	Head of Intellectual Information Technologies Department of BSUIR
Prof.Dr.	Vladimir Golovko	Head of Intellectual Information Technologies Department of BrSTU
Prof.Dr.	Alexander Ivaniuk	Professor of Computer Science Department of BSUIR
Ph.D.	Dmitry Kukin	Head of Computational Methods and Programming Department of BSUIR
Ph.D.	Alexander Markov	Head of Control Systems Department of BSUIR
Ph.D.	Anatoly Navrotsky	Head of Information Technologies in Automated Systems Department of BSUIR
Ph.D.	Maxim Davydov	Head of Theoretical Bases of Electrical Engineering Department of BSUIR
Prof.Dr.	Stanislav Kolosov	Professor of Computational Methods and Programming Department of BSUIR

Пленарные доклады

Plenary session

НОВЫЕ КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ, АНАЛИЗ И ИНТЕРПРЕТАЦИЯ ГЕОДАННЫХ

Таранчук В. Б.

Кафедра компьютерных технологий и систем, Факультет прикладной математики и информатики,

Белорусский государственный университет

Минск, Республика Беларусь

E-mail: taranchuk@bsu.by

Обсуждаются методические аспекты и технические решения вопросов разработки и сопровождения геологических, геоэкологических моделей на основе интеграции современных географических информационных систем и систем компьютерной алгебры. На примере разработанной программной системы создания и анализа геологической модели демонстрируются соответствующие средства, возможности комбинировать инструменты географических информационных систем с модулями статистического анализа данных, математически сложных вычислений, включающих реализации современных методов и алгоритмов анализа и интерпретации пространственных данных, нейросетевого аппарата.

ВВЕДЕНИЕ

Построение компьютерных геологических, геоэкологических моделей в настоящее время является обязательной частью решений экспертов во многих областях. В частности, при мониторинге состояния окружающей среды, недр, решении проблем рационального использования минерально-сырьевой базы, в проектах о защитных мерах, связанных с описанием рельефа и инженерно-геологического строения местности, при планировании контроля и управления характеристиками подземных участков, существующей подземной инфраструктуры. Геологические модели имеют особую роль для обоснования проектирования и строительства подземных сооружений, и будут всегда иметь решающее значение при создании подземных сооружений с более низкими первоначальными издержками и рисками, продолжительным жизненным циклом.

I. ОСНОВЫ ГЕОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ.

В текущем понимании геологическая модель содержит обобщенное описание состава, структуры, формы объектов, состояния изучаемого фрагмента земной коры, поверхности на разных стадиях. Она включает не только параметры геологического строения, но и цифровые характеристики, обеспечивающие получение и графическую визуализацию объемного распределения структурно-вещественных комплексов и компонент модели. На современном этапе развития геологической науки и компьютерных технологий есть возможность комплексирования всей имеющейся геолого-геофизической и сопутствующей информации, ее интегрированного анализа с помощью вычислительных экспериментов. В то же время, многие специалисты отмечают, что понятие геологической модели до настоящего времени не в полной мере формализовано, созда-

ние компьютерных геологических моделей требует унификации ряда процедур.

Базой геологического моделирования, включающего структурную геологию, седиментологию, стратиграфию, диагенез, являются методы прикладной математики, алгоритмы обработки данных, реализующие их компьютерные программы. Соответствующее программное обеспечение компьютерных геологических моделей включает загрузку из разных источников и предобработку данных, корреляцию, формирование цифровых кубов фильтрационно-емкостных свойств, интерактивный анализ данных, визуализацию с применением графики разных типов, картопостроение.

II. КОМПЬЮТЕРНЫЕ ГЕОЛОГИЧЕСКИЕ МОДЕЛИ И ИХ ОСОБЕННОСТИ.

Создание цифровых геологических, геоэкологических моделей является относительно молодым направлением, оно активно развивается. При этом многие специалисты отмечают, что вопросы методики и технологии построения геологических моделей остаются очень сложной инженерной задачей, над решениями работают многие специалисты, в том числе в Стэнфордском университете, Норвежском компьютерном центре, Французском институте нефти, Научной школе в Нанси. Применение компьютерных геологических моделей расширяется, но пока в странах СНГ крупные государственные и частные компании потребители соответствующих программных комплексов предпочитают системы мировых лидеров. Актуальными являются вопросы разработки и внедрения альтернативных компьютерных геологических моделей. Важнейшими составляющими при этом являются: простота освоения, многофункциональность, возможности взаимообмена результатами при использовании нескольких программных продуктов, наличие средств оценки адекватности и точности получаемых цифровых моделей. Ключевыми яв-

ляются вопросы систематизации исходной и получаемой в вычислительных экспериментах информации, формирования баз знаний, автоматизации настройки и адаптации моделей, их сопоставления и оценки точности.

Моделирование объектов геологии предполагает использование сложных математических методов. Программирование, реализация алгоритмов построения и адаптации геологических, геоэкологических моделей трудоемко. Из опыта разработки, сопровождения и внедрения программных комплексов «ГеоБазаДанных» ([1, 2]), «ГБД-п», «ГБД-э» ([3, 4]) следует, что для поддержания подобных программных продуктов в актуальном состоянии, отвечающем постоянно возрастающим аппаратным возможностям, нужен большой коллектив опытных программистов. Учитывая текущие возможности по обеспеченности системами компьютерной математики, более эффективным, с заметно более низкими трудозатратами на разработку и сопровождение, относительно быстрым по времени будет подход, основанный на интеграции в единый программный комплекс модулей современных версий систем компьютерной алгебры (СКА) и географических информационных систем (ГИС). При этом следует понимать, что для решения задач обработки исходных данных, в частности, результатов дистанционного зондирования, сейсмо- и магниторазведки, моделирования, никакая конкретная ГИС не является полным набором пространственно-аналитических методов и средств анализа. Во многих случаях приходится комбинировать инструменты ГИС с программами статистического анализа данных, средствами для математически сложных вычислений, включающих реализации современных методов и алгоритмов анализа и интерпретации пространственных данных.

Теоретической базой при решении задач математического моделирования объектов геологии, геоэкологии является концепция, следуя которой ядром и основой для построения компьютерных геологических моделей является цифровое описание ограничивающих объём поверхностей – структурный каркас и набор используемых литологических типов. При этом основным этапом является построение генерализованных поверхностей, описывающих топологию объекта, последовательность залегания геологических тел, слоев, то есть своеобразная структурная «этажерка» [5, 2, 4]. Для структурированной по слоям трехмерной геологической модели применим подход построения в режиме «конструктор», когда сборка и редактирование модели производятся по частям, которыми служат отдельные геологические элементы. Для слоев в описание включаются распределения изучаемых параметров. Исходными данными для этих описаний, как правило, являются значения на-

блюдаемого параметра в пунктах контроля, которые размещены на площади нерегулярно. Например, фиксируются, контролируются характеристики: стрессдеформация, абразивность, теплопроводность, пористость, проницаемость, гидропроводность, описания геологических изменений и истории развития участка – соответствующие замеры могут рассматриваться как параметры, их распределение в объеме включаются в цифровые кубы данных.

III. КОМПЬЮТЕРНЫЕ ГЕОЛОГИЧЕСКИЕ МОДЕЛИ И ИХ ПЕРСПЕКТИВЫ.

Представляется, что на данном этапе основной разработки и сопровождения компьютерных геологических моделей должны быть новые «умные» методы адаптации моделей в процессе их эксплуатации, интеллектуальные вычислительные системы, реализующие методы «самонастройки» моделей с учётом дополняемых данных фактического развития процессов. Элементы такого подхода реализованы в компьютерном программном комплексе «Генератор геологической модели залежи» (ГГМЗ). Основные функциональные возможности, подтверждающие эффективность предлагаемого подхода соответствующие программные компоненты ГГМЗ, примеры для разных вариантов анализа и интерпретаций можно посмотреть в [6 – 8].

В докладе с использованием инструментов комплекса ГГМЗ будут показаны представительные примеры, иллюстрирующие возможности графической визуализации моделей, традиционных методов предобработки первичной информации, применения при интерпертации исходных данных и результатов моделирования нейронных сетей, многоточечной статистики.

Отдельное внимание будет уделено демонстрации возможностей графической визуализации и сопоставления результатов средствами системы компьютерной математики Wolfram *Mathematica* ([9, 10]), интеграции получаемых карт и иллюстраций в географические информационные системы. Целесообразность подобного анализа обусловлена желанием показать приемы получения эталонов и обоснования экспертных решений, которые с учетом обрабатываемой геологической информации всегда субъективны.

Возможности системы *Mathematica* по статистической обработке данных ([11]) будут отмечены в ракурсе относительно простого включения в число инструментов анализа дополнительных математически сложных алгоритмов анализа многомерных данных [12]. Подобное очень важно, потому что методы анализа типа цитированных могут найти применение в практике только при наличии соответствующих программных модулей – используя СКА *Mathematica*, не имеющие специальной математической подготовки пользователи, могут освоить и приме-

нять сложные алгоритмы, потому что в системе есть полные и содержательные пояснения реализованных алгоритмов и программных функций.

Обсуждение и уточнение возможных дополнительных направлений развития аппарата нейронных сетей [13, 14], их применения при создании интеллектуальных программных средств анализа и конструирования геологических моделей с реализацией идеи «сеть сама учится выполнять задачу» предполагается инициировать демонстрацией в реализованных по методике [15] модулях комплекса специально подобранных примеров.

ЗАКЛЮЧЕНИЕ

Рассмотрены вопросы, сформулированы конкретные предложения проектирования и разработки программных систем для создания и сопровождения геологических, геоэкологических моделей путем интеграции в единый программный комплекс модулей современных версий систем компьютерной алгебры и географических информационных систем.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Буйкис, А. А. О программном обеспечении вычислительного эксперимента в задачах неизоэотермической двухфазной многокомпонентной фильтрации / А. А. Буйкис, А. Б. Золотухин, В. Б. Таранчук // Прикладные задачи математической физики. Рига, ЛГУ. –1985. –С. 191–205.
2. Барвенов, С. А. Методика, инструментарий адаптации математических моделей процессов подземной гидродинамики / С. А. Барвенов, М. Ф. Кибаш, В. Б. Таранчук // Выбранные научные работы БДУ, т. «Математика», Минск. –2001. –С. 34–65.
3. Таранчук, В. Б. Интерактивная компьютерная система визуализации процессов переноса примесей в почвогрунтах «ГБД-п» / В. Б. Таранчук, В. В. Таранчук // Природопользование. –2005. –Вып. 11. –С. 139–143.
4. Таранчук, В. Б. Построение цифровых геолого-экологических моделей в системе ГБД-э / Геоэкология Беларуси / М. Г. Ясовеев [и др.] ; под ред. М. Г. Ясовеева, В. Б. Таранчука. – Минск : Право и экономика. –2006. Научное издание. –С. 72–96.
5. Булыгин, Д. В. Геология и имитация разработки залежей нефти / Д. В. Булыгин, В. Я. Булыгин. – М. : Недра. –1996. – 382 с.
6. Таранчук, В.Б. Программный комплекс адаптации геологических моделей. Концепция, решения, примеры реализации / В. Б. Таранчук // Проблемы физики, математики и техники. –2017. –№ 3 (32). –С. 81–90.
7. Таранчук, В. Б. Интегрированный программный комплекс тестировщика геологических моделей / В. Б. Таранчук // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. –2017. –№ 16 (265). –С. 148–159.
8. Taranchuk, V. B. The integrated computer complex of an estimation and adapting of digital geological models / V. B. Taranchuk // Studia i Materiały. Volume 14. –2017. –№ 2. –С. 73–86 (in Engl.)
9. Таранчук, В. Б. Особенности функционального программирования интерактивных графических приложений / В. Б. Таранчук // Вестник Самарского государственного университета. Естественная серия, раздел Математика: –2015. –№ 6 (128). –С. 178–189.
10. Taranchuk, V. B. Methods and tools of development of interactive teaching materials on the computer mechanics / V. B. Taranchuk, M. A. Zhuravkov // Studia i Materiały. –2016. –№ 11. –С. 69–83 (in Engl.)
11. Таранчук, В. Б. О возможностях и инструментах реализации вероятностно-статистических методов в системе Mathematica / В. Б. Таранчук // Теория вероятностей, случайные процессы, математическая статистика и приложения : материалы Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г. А. Медведева, Минск : РИВШ . –2015. –С. 320–325.
12. Муха, В. С. Математические модели многомерных данных / В. С. Муха // Доклады БГУИР. –№ 2 (80). –2014. –С. 143–158.
13. Golenkov, V. V. From training intelligent systems to training their development tools / V. V. Golenkov, N. A. Gulyakina, N. V. Grakova, I. T. Davydenko, V. Y. Nikulenko, A. P. Egemeev, V. B. Tarasov / Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2018) : материалы Междунар. научн.- техн.конф. Мн. : БГУИР. –2018. –С. 81–98
14. Golovko, V. A. Integration of artificial neural networks and knowledge bases / V. A. Golovko, V. V. Golenkov, V. P. Ivashenko, V. V. Taberko, D. S. Ivaniuk, A. A. Kroshchanka, M. V. Kovalev / Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2018) : материалы Междунар. научн.- техн.конф. Мн. : БГУИР. –2018. –С. 133–145
15. Таранчук, В. Б. Возможности и средства Wolfram Mathematica для разработки интеллектуальных обучающих систем / В. Б. Таранчук // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика: –2015. –№ 1 (198) том 33. –С. 102–110.

Основываясь на наибольшей в мире коллекции специальных функций и двух десятилетиях развития символьных и численных алгоритмов, система *Mathematica*, начиная с версии 8, предлагает беспрецедентный уровень поддержки параметрических распределений [7].

ТЕКУЩЕЕ СОСТОЯНИЕ И НАПРАВЛЕНИЯ РАЗВИТИЯ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Голенков В. В., Гулякина Н. А., Шункевич Д. В.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: golen@bsuir.by, guliakina@bsuir.by, shunkevich@bsuir.by

В работе рассмотрены основные актуальные проблемы в области разработки современных интеллектуальных систем. Предложен подход к их решению, основанный на использовании открытой семантической технологии проектирования интеллектуальных систем

ВВЕДЕНИЕ

До настоящего времени традиционные информационные технологии и технологии искусственного интеллекта развивались независимо друг от друга.

Современное состояние информационных технологий в целом можно охарактеризовать как

- иллюзию благополучия;
- иллюзию всеислия финансовых ресурсов в решении различных технических задач;
- «вавилонское столпотворение» различных технических решений, о совместимости которых никто серьезно не задумывается;
- отсутствие комплексного системного подхода к автоматизации сложных видов деятельности;
- отсутствие осознания того, что недостатки современных информационных технологий имеют фундаментальный, системный характер.

Сейчас настало время фундаментального переосмысления опыта эволюции традиционных информационных технологий и их глубокой интеграции с технологиями искусственного интеллекта. Это необходимо для устранения целого ряда недостатков традиционных информационных технологий.

I. НЕДОСТАТКИ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

К недостаткам современных информационных технологий можно отнести:

1. Многообразие семантически эквивалентных форм (языков) представления (кодирования) обрабатываемой информации (знаний) в памяти компьютерных систем. Следствием указанного многообразия является многообразие семантически эквивалентных форм представления в памяти компьютерных систем способов решения задач (как процедурных, так и непроцедурных – функциональных, логических и т.д.). Отсутствие унификации представления различного вида знаний в памяти современных компьютерных систем приводит:

- к дублированию семантически эквивалентных информационных компонентов компьютерных систем;
 - к семантической несовместимости компьютерных систем и, следовательно, к высокой трудоемкости их интеграции в системы более высокого уровня иерархии;
 - к существенному снижению эффективности применения методики компонентного проектирования компьютерных систем на основе библиотеки многократно используемых компонентов (особенно, если речь идет о «крупных» компонентах, в частности, о типовых подсистемах).
2. Недостаточно высокую степень обучаемости современных компьютерных систем в ходе их эксплуатации, следствием чего является высокая трудоемкость их сопровождения и совершенствования, а также недостаточно длительный их жизненный цикл. Обучаемость компьютерной системы обеспечивается
 - гибкостью – многообразием и трудоемкостью возможных изменений, вносимых в систему в процессе пополнения системы новыми знаниями и навыками и совершенствования уже приобретенных знаний и навыков;
 - стратифицированностью системы – четким разделением системы на достаточно независимые друг от друга уровни иерархии;
 - рефлексивностью – способностью анализировать собственное состояние и свою деятельность;
 - гибридность – способностью приобретать и использовать широкое (а в идеале – неограниченное) многообразие знаний и навыков;
 - уровнем самообучаемости – уровнем автоматизации приобретения новых знаний и навыков, а также совершенствования уже приобретенных знаний и навыков

3. Отсутствие возможности у экспертов реально влиять на качество разрабатываемых компьютерных систем. Опыт разработки сложных компьютерных систем показывает, что посредничество программистов между экспертами и проектируемыми компьютерными системами существенно искажает вклад экспертов. При разработке компьютерных систем следующего поколения доминировать должны не программисты, а эксперты, способные точно излагать свои знания.
4. Отсутствие семантической (смысловой) унификации интерфейсной деятельности пользователей компьютерных систем, что вместе с многообразием форм реализации пользовательских интерфейсов приводит:
 - к серьезным накладным расходам на усвоение пользовательских интерфейсов новых компьютерных систем;
 - к неполному и неэффективному использованию возможностей эксплуатируемых компьютерных систем.

Преодолеть указанные недостатки можно только путем фундаментального переосмысления архитектуры и принципов организации сложных компьютерных систем. Основой такого переосмысления является устранение многообразия форм представления (кодирования) информации в памяти компьютерных систем.

Преодоление недостатков современных компьютерных систем предполагает:

- унификацию представления обрабатываемой информации
- функциональную унификацию (унификацию принципов обработки информации)

Важно отметить, что унификация (стандартизация) – это не ограничение творческой свободы инженера, а способ обеспечения совместимости результатов инженерной деятельности.

II. ПРИНЦИПЫ УНИФИКАЦИИ ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ

Объективным ориентиром унификации представления информации в памяти компьютерных систем является формализация смысла представляемой информации.

Уточнение смыслового представления информации основано на максимально возможном упрощении синтаксиса при обеспечении универсальности.

Унификация информации, используемой в компьютерных системах, предполагает:

- синтаксическую унификацию используемой информации – унификацию формы представления (кодирования) этой информации. При этом следует отличать
 - кодирование информации в памяти компьютерной системы;

- представление информации для людей, обеспечивающее однозначность интерпретации (понимания, трактовки) этой информации разными людьми;
- семантическую унификацию используемой информации в основе которой лежит согласование и точная спецификация всех (!) используемых понятий (концептов) с помощью иерархической системы формальных онтологий.

Критерием эффективности синтаксической унификации представления информации любого вида в памяти компьютерных систем естественно считать максимально возможное упрощение синтаксиса путем исключения из такого внутреннего универсального языка всех средств, обеспечивающих коммуникационную функцию языка (т. е. обмен сообщениями). Для внутреннего языка компьютерной системы являются излишними коммуникационные возможности, такие как имена, союзы, предлоги, разделители, ограничители, склонения, спряжения и т. д.

Указанный внутренний универсальный язык компьютерных систем фактически есть не что иное, как язык смыслового представления знаний. Очевидно, что унификация формы представления информации в памяти компьютерных систем на основе формального представления ее смысла выглядит вполне логично, т. к. только смысл информации является единственной объективной основой для унификации ее представления (кодирования). Таким образом, формализация смысла – ключ к решению многих проблем разработки современных компьютерных систем.

Согласно В. В. Мартынову, «фактически всякая мыслительная деятельность человека (не только научная), как полагают многие ученые, использует внутренний семантический код, на который переводят с естественного языка и с которого переводят на естественный язык. Поразительная способность человека к идентификации огромного множества структурно различных фраз с одинаковым смыслом и способность запомнить смысл вне этих фраз убеждает нас в этом» [1].

Смысл – это абстрактная знаковая конструкция, являющаяся инвариантом максимального класса семантически эквивалентных знаковых конструкций (текстов), принадлежащих разным языкам и удовлетворяющая следующим требованиям:

- отсутствие синонимии знаков (многократного вхождения знаков с одинаковыми денотатами);
- отсутствие дублирования информации в виде семантически эквивалентных текстов (не путать с логической эквивалентностью);

- отсутствие омонимичных знаков (в том числе местоимений);
- отсутствие у знаков внутренней структуры (атомарный характер знаков);
- отсутствие склонений, спряжений (как следствие отсутствия у знаков внутренней структуры);
- отсутствие фрагментов знаковой конструкции, не являющихся знаками (разделителей, ограничителей, и т.д.);
- выделение знаков связей, компонентами которых могут быть любые знаки, с которыми знаки связей связываются синтаксически задаваемыми отношениями инцидентности.

Существенно подчеркнуть, что рассмотренные принципы смыслового представления информации приводят к нелинейным (графовым) знаковым конструкциям, что в некотором смысле усложняет организацию памяти компьютерных систем, но существенно упрощает их логическую архитектуру.

III. УНИФИКАЦИЯ ПРИНЦИПОВ ОБРАБОТКИ ИНФОРМАЦИИ

В качестве основы для унификации принципов обработки информации предлагается использовать многоагентный подход, который в настоящее время широко используется в различных областях. Ориентация на многоагентный подход как основу для организации обработки информации обусловлена следующими основными преимуществами такого подхода [2, 3]:

- автономность (независимость) агентов в рамках такой системы, что позволяет локализовать изменения, вносимые в систему при ее эволюции, и снизить соответствующие трудозатраты;
- децентрализация обработки, т.е. отсутствие единого контролирующего центра, что также позволяет локализовать вносимые в многоагентную систему изменения.

Однако, к основным недостаткам большинства популярных современных средств построения многоагентных систем можно отнести следующие:

- большинство современных средств построения многоагентных систем ориентированы на представление знаний агента при помощи узкоспециализированных языков, зачастую не предназначенных для представления знаний в широком смысле. Речь при этом идет как о знаниях агента о себе самом, так и о знаниях о внешней среде. В некоторых подходах вначале строится онтология, для создания которой, однако, часто используются средства с низкой выразительной способностью, не предназначенные для построения онтологий. В конечном итоге такой подход приводит к сильной

ограниченности возможностей разработанных многоагентных систем и их несовместимости;

- абсолютное большинство современных средств предполагает, что взаимодействие агентов осуществляется путем обмена сообщениями непосредственно от агента к агенту. Такой подход обладает существенным недостатком, связанным с тем, что в этом случае каждый агент системы должен иметь достаточно полную информацию о других агентах в системе, что приводит к дополнительным затратам ресурсов, кроме того добавление или удаление одного или нескольких агентов приводит к необходимости оповещения об этом других агентов;
- многие средства построения многоагентных систем построены таким образом, что логический уровень взаимодействия агентов жестко привязан к физическому уровню реализации многоагентной системы. Например, при передаче сообщений от агента к агенту разработчику многоагентной системы необходимо помимо семантически значимой информации указывать IP-адрес компьютера, на котором расположен агент-получатель, кодировку, с помощью которой закодирован текст сообщения и другую техническую информацию, обусловленную исключительно особенностями текущей реализации средств;
- в большинстве подходов среда, с которой взаимодействуют агенты, уточняется отдельно разработчиком для каждой многоагентной системы, что с одной стороны, расширяет возможности применения соответствующих средств, но с другой стороны приводит к существенным накладным расходам и несовместимости таких многоагентных систем. Кроме того, в ряде случаев разработчик также обязан учитывать особенности технической реализации средств разработки в плане их стыковки с предполагаемой средой, в роли которой может выступать, например, локальная или глобальная сеть.

В рамках данной работы перечисленные недостатки предполагается устранять за счет использования следующих принципов:

- коммуникацию агентов предлагается осуществлять по принципу «доски объявлений», однако, в отличие от классического подхода, в роли сообщений выступают спецификации в общей памяти компьютерной системы выполняемых агентами действий (процессов), направленных на решение каких-либо задач, а в роли среды коммуникации выступает сама эта память. Такой подход позволяет:

- исключить необходимость разработки специализированного языка для обмена сообщениями;
- обеспечить «обезличенность» общения, т.е. каждый из агентов в общем случае не знает, какие еще агенты есть в системе, кем сформулирован и кому адресован тот или иной запрос. Таким образом, добавление или удаление агентов в систему не приводит к изменениям в других агентах, что обеспечивает модифицируемость всей системы;
- агенты, в том числе, конечный пользователь, получают возможность формулировать задачи в *декларативном ключе*, т.е. не указывать для каждой задачи способ ее решения. Таким образом, агенту заранее не нужно знать, каким образом система решит ту или иную задачу, достаточно лишь специфицировать конечный результат;
- в роли внешней среды для агентов выступает та же общая память, в которой формулируются задачи и посредством которой осуществляется взаимодействие агентов. Такой подход обеспечивает унификацию среды для различных систем агентов, что, в свою очередь, обеспечивает их совместимость.
- спецификация каждого агента описывается средствами языка представления знаний в той же памяти, в которой осуществляется решение задач, что позволяет:
 - минимизировать число специализированных средств, необходимых для спецификации агентов, как языковых, так и инструментальных;
 - с одной стороны - минимизировать необходимую в общем случае спецификацию агента, которая включает условие его инициирования и программу, описывающую алгоритм работы агента, с другой стороны - обеспечить возможность неограниченного расширения такой спецификации для каждого конкретного случая;
- синхронизацию деятельности агентов предполагается осуществлять на уровне выполняемых ими процессов, направленных на решения тех или иных задач. Таким образом, каждый агент трактуется как некий абстрактный процессор, способный решать задачи определенного класса.
- каждый информационный процесс в любой момент времени имеет ассоциативный доступ к необходимым фрагментам базы знаний, хранящейся в общей памяти, за исключением фрагментов, заблокированных другими процессами в соответствии с исполь-

зуемым механизмом синхронизации выполнения параллельных процессов. Таким образом, с одной стороны, исключается необходимость хранения каждым агентом информации о внешней среде, с другой стороны, каждый агент, как и в классических многоагентных системах, обладает только частью всей информации, необходимой для решения задачи.

Важно отметить, что в общем случае невозможно априори предсказать, какие именно знания, модели и способы решения задач понадобятся системе для решения конкретной задачи. В связи с этим необходимо обеспечить, с одной стороны, возможность доступа ко всем необходимым фрагментам базы знаний (в пределе - ко всей базе знаний), с другой стороны - иметь возможность локализовать область поиска пути решения задачи, например, рамками одной предметной области.

IV. ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТРЕБОВАНИЯ К ИХ РАЗРАБОТЧИКАМ

Таким образом, устранение проблем современных информационных технологий фактически преобразует современные компьютерные системы в интеллектуальные системы. Следовательно, интеллектуальные системы – это не альтернативная ветвь развития компьютерных систем, а естественный этап их эволюции, направленной на повышение уровня гибкости, стратифицированности, рефлексивности, гибридности, и обучаемости компьютерных систем.

С учетом сказанного, можно определить *интеллектуальную систему* как компьютерную систему, основанную на знаниях, обладающую высоким уровнем обучаемости, т.е. способностью быстро приобретать новые и совершенствовать уже приобретенные знания и навыки и при этом не иметь никаких ограничений на вид приобретаемых и совершенствуемых ею знаний и навыков, а также на способ их совместного использования.

Важно подчеркнуть, что переход от проектирования традиционных компьютерных систем к проектированию интеллектуальных систем требует от разработчиков существенного повышения:

- уровня математической культуры при построении формальной модели среды, в которой функционирует интеллектуальная система, формальных моделей решаемых ею задач и формальных моделей различных используемых ею способов решения задач;
- уровня системной культуры, позволяющей адекватно оценивать качество разрабатываемых систем с точки зрения общей теории систем и, в частности, оценивать об-

щий уровень автоматизации, реализуемый с помощью этих систем;

- уровня технологической культуры, обеспечивающей совместимость разрабатываемых систем, а также постоянное расширение библиотеки многократно используемых компонентов создаваемых систем.

V. РЕАЛИЗАЦИЯ ПРЕДЛОЖЕННЫХ ПРИНЦИПОВ УНИФИКАЦИИ ПРЕДСТАВЛЕНИЯ И ОБРАБОТКИ ИНФОРМАЦИИ

Рассмотренные принципы унификации представления и обработки информации реализуются в виде открытой семантической технологии компонентного проектирования гибких совместимых гибридных компьютерных систем нового поколения, названной нами Технологией OSTIS (Open Semantic Technology for Intelligent Systems) [4, 5].

Перечислим основные принципы предлагаемой семантической технологии:

- ориентация на смысловое однозначное представление знаний в виде семантических сетей, имеющих базовую теоретико-множественную интерпретацию, что обеспечивает решение проблемы многообразия форм представления одного и того же смысла, и проблемы неоднозначности семантической интерпретации информационных конструкций;
- использование ассоциативной графодинамической памяти;
- применение агентно-ориентированной модели обработки знаний;
- реализация предлагаемой технологии в виде интеллектуальной Метасистемы IMS.ostis [6], которая построена по этой же технологии и осуществляет поддержку проектирования компьютерных систем, разрабатываемых по указанной технологии;
- обеспечение в проектируемых системах высокого уровня гибкости, стратифицированности, рефлексивности, гибридности и, как следствие, обучаемости.

В основе предлагаемой семантической технологии лежит базовый унифицированный язык смыслового кодирования информации в памяти компьютерных систем, удовлетворяющий перечисленным ранее требованиям и названный **SC-кодом** (Semantic Computer Code). Тексты SC-кода (sc-тексты) представляют собой семантические сети с базовой теоретико-множественной интерпретацией. Элементы таких сетей называются *sc-элементами* (sc-узлами, sc-дугами, sc-ребрами). Теоретико-множественная интерпретация предполагает, что каждый sc-элемент может быть либо знаком множества sc-элементов, либо знаком терминальной сущности (сущности, не являющейся множеством). В свою очередь,

sc-дуги и sc-ребра трактуются как знаки двухмощных множеств (ориентированных или неориентированных соответственно), элементами которых являются sc-элементы, инцидентные соответствующей sc-дуге или sc-ребру. При этом выделяется класс *sc-дуг основного вида*, каждая из которых трактуется как принадлежность элемента множеству. Таким образом, SC-код как язык семантических сетей объединяет два базовых аспекта представления информации – синтаксический (выделение синтаксических типов sc-элементов – sc-дуг, sc-узлов и т.д.), и семантический (выделение базовой типологии сущностей, обозначаемых sc-элементами, – терминальных сущностей, множеств, отношений и т.д.).

Системы, построенные на основе Технологии OSTIS, называются ostis-системами. Каждая ostis-система в общем случае имеет архитектуру, представленную на рисунке 1.

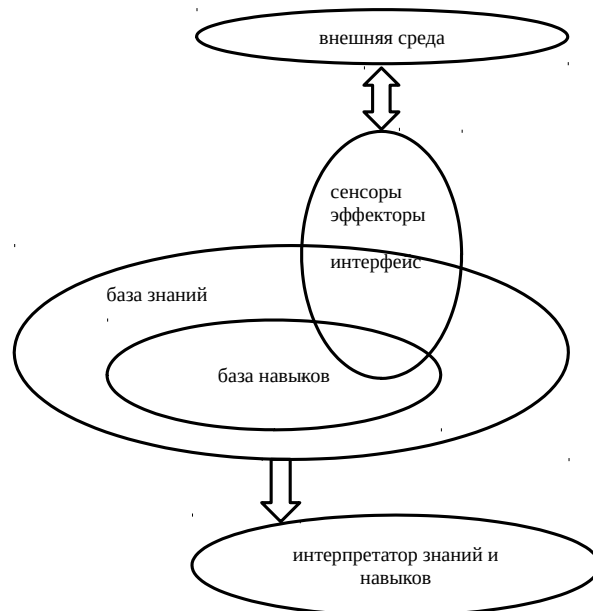


Рис. 1 – Архитектура ostis-системы

Под **базой знаний** ostis-системы понимается систематизированная совокупность всех знаний и навыков, хранимых в памяти такой системы. Как видно из рисунка, частью базы знаний является **база навыков**, то есть систематизированный комплекс всех известных интеллектуальной системе способов решения различных задач и классов задач. Возможность описания в базе знаний способов решения различных задач достигается за счет:

- наличия в технологии формальных средств описания в базе знаний алгоритмов, программ любого уровня сложности (как императивных, так и декларативных);
- описания в базе знаний системы спецификаций всех агентов обработки знаний;
- описания в базе знаний системы программ агентов обработки знаний, что становится

возможным благодаря наличию в рамках технологии графового процедурного языка программирования SCP (Semantic Code Programming), который, с одной стороны, ориентирован на обработки sc-текстов, а с другой стороны, тексты программ которого записываются в SC-коде, т.е. являются sc-текстами.

Таким образом, каждая ostis-система состоит из полной модели этой системы, описанной средствами SC-кода и хранящейся в базе знаний (sc-модели ostis-системы) и интерпретатора знаний и навыков (платформы интерпретации sc-моделей ostis-систем), который в общем случае может быть реализован как программно, так и аппаратно. При этом если фундаментальной основой традиционных информационных технологий и соответствующих ей компьютерных архитектур является абстрактная машина фон Неймана, то основой ostis-систем как нового этапа эволюции компьютерных систем должны стать семантические параллельные компьютеры с нелинейной ассоциативной реконфигурируемой памятью. Это очевидное направление повышения производительности компьютерных систем нового поколения.

Интерпретатор знаний и навыков включает такие обязательные компоненты как хранилище sc-текстов (sc-хранилище), обеспечивающее хранение базы знаний и доступ к ней, а также интерпретатор программ языка SCP (scp-программ).

Приведенная архитектура позволяет обеспечить платформенную независимость ostis-систем, в том числе возможность легкого перевода разработанной системы с программной платформы на аппаратную.

Интерфейс ostis-системы в рамках предлагаемого подхода трактуется как подсистема, которая строится по тем же принципам, то есть имеет свою специализированную базу знаний и базу навыков.

Достоинства предлагаемой семантической технологии заключаются в том, что она:

- ориентирована на разработку компьютерных систем нового поколения (гибридных и семантических совместимых компьютерных систем с высокой степенью обучаемости);
- имеет открытый характер как для ее пользователей (разработчиков прикладных интеллектуальных систем), так и для тех, кто желает участвовать в ее совершенствовании;
- ориентирована на постоянное повышение темпов ее эволюции;
- является основой для решения проблем семантической совместимости самых раз-

личных научных и технических знаний, так как она ориентирована на формализацию междисциплинарных связей самого различного вида.

ЗАКЛЮЧЕНИЕ

В работе рассмотрены основные недостатки современных компьютерных систем, предложены подходы к их устранению, реализуемые в виде открытой семантической технологии проектирования интеллектуальных систем.

Следует подчеркнуть, что исследования в области искусственного интеллекта носят ярко выраженный междисциплинарный характер. О необходимости глубокой конвергенции различных научных направлений в области искусственного интеллекта в целях построения общей формальной теории интеллектуальных систем говорится в целом ряде работ, в частности, Янковской А.Е. [7] и Палагина А.В. [8]. Этому также посвящены конференции «Artificial General Intelligence», проводимые с 2008 года [9]. Основой для построения такой теории могут стать предлагаемые в работе принципы унификации представления и обработки информации и построенная на их основе технология проектирования интеллектуальных систем.

СПИСОК ЛИТЕРАТУРЫ

1. Мартынов, В.В. Семиологические основы информатики / В.В. Мартынов - Мн.: Наука и техника, 1974. - 192 с.
2. Wooldridge, M. An introduction to multiagent systems / M. Wooldridge. - 2nd ed. - Chichester : J. Wiley, 2009. - 484 p.
3. Тарасов, В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В. Б. Тарасов. - М. : Эдиториал УРСС, 2002. - 348 с.
4. Голенков, В. В. Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 1: Принципы создания / В. В. Голенков, Н. А. Гулякина // Онтология проектирования. - 2014. - № 1. - С. 42-64.
5. Голенков, В. В. Проект открытой семантической технологии компонентного проектирования интеллектуальных систем. Часть 2: Унифицированные модели проектирования / В. В. Голенков, Н. А. Гулякина // Онтология проектирования. - 2014. - № 4. - С. 34-53.
6. Метасистема IMS.ostis [Электронный ресурс]. - Режим доступа: <http://ims.ostis.net>. - Дата доступа: 10.09.2018.
7. Янковская, А.Е. Анализ данных и знаний на основе конвергенции нескольких наук и научных направлений / А.Е. Янковская // Международная конференция «Интеллектуализация обработки информации» (ИОИ-8), Кипр, г. Пафос, 17-24 октября 2010 г. - сс. 196-199.
8. Палагин, А.В. Проблемы трансдисциплинарности и роль информатики / А.В. Палагин // Кибернетика и системный анализ. - 2013. №5. - сс. 3-13.
9. Конференция Artificial General Intelligence (AGI) <http://agi-conference.org>

О НЕКОТОРЫХ ВОПРОСАХ АЛГОРИТМИЧЕСКОГО МАРКЕТИНГА

Пархименко В. А.

Кафедра экономики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: parkhimenko@bsuir.by

Рассматривается история становления алгоритмического маркетинга. Раскрывается сущность и основные задачи алгоритмического маркетинга.

I. СТАНОВЛЕНИЕ АЛГОРИТМИЧЕСКОГО МАРКЕТИНГА

В науке является естественным и даже обязательным ориентироваться на строгую формализацию предметной области, поэтому даже маркетинг, который осторожно определяют в учебниках как науку и искусство (т.е. сложное сочетание интуиции, опыта, врожденных способностей и здравого смысла), всегда был объектом многочисленных попыток математического моделирования и перевода в однозначные алгоритмы принятия решений.

Так, уже в конце 1950-х годов стали появляться отдельные публикации об использовании математических моделей в маркетинге, но, по всей видимости, отправной точкой в этом контексте следует считать публикацию классической монографии Филипа Котлера «Принятие маркетинговых решений: модельный подход» [1], изданной в 1971 году и охватывающей основные существующие на тот момент математические модели в маркетинге. В монографии Ф. Котлер делает очевидное (однако на тот момент сугубо теоретического характера) утверждение, что, имея в распоряжении математическую модель и набор ограничений, можно поручить компьютеру поиск решения, оптимизирующего целевую функцию, определяемую владельцем бизнеса.

Монография переиздавалась с существенными дополнениями и изменениями в 1983 [2] и в 1992 [3] годах, при этом Ф. Котлер, став к тому времени признанным "отцом-основателем" современного маркетинга, уступил место ведущего автора Гэри Лилиэну.

Последний в дальнейшем (в конце 1990-х – начале 2000-х гг.) создал свое собственное новое направление в маркетинговой науке – маркетинговую инженерию (marketing engineering), т.е. методологию построения и использования математических моделей и компьютерных технологий для принятия научно-обоснованных решений в маркетинговой деятельности. Г. Лилиэн организовал консалтинговую компанию "DecisionPro", разработал специализированное программное обеспечение и издал ряд учебников по маркетинговой инженерии в 2004 г. [4] и в 2017 г. [5].

Начиная с 1989 года неоднократно переиздавалась еще одна классическая монография «Модели реакции рынка: эконометрический анализ и временные ряды» [6]. В издании 2006 года авторы впервые отметили, что разрабатываемая десятилетиями теория и математические модели стали наконец-то активно использоваться на практике. Фактором этого послужило бурное развитие информационных технологий, позволивших накапливать большие объемы данных о покупательском поведении с дальнейшим их анализом и выходом на оптимизационные задачи.

Серьезными обзорными работами в последние десятилетия были два издания (2008, 2017 гг.) коллективной монографии «Справочник по моделям принятия маркетинговых решений» под редакцией Б. Виеренга и Р. Ван дер Ланса [7, 8]. Отметим, что авторы также указывают на ведущую роль информационных технологий в прогрессе своей области научных исследований, что, помимо прочего, привело к появлению новых 13 глав (из 17!) во втором издании монографии, т.е. всего за 9 лет произошли существенные изменения.

В Беларуси математическими моделями в маркетинге серьезно занимались и опубликовали обзорные книги Н.Н. Анохина [11], А.В. Сак и В.А. Журавлев [12]. (Стоит при этом отметить, что В.А. Журавлев – инициатор создания, разработчик типовой учебной программы и бессменный лектор с 2013 года двухсеместровой учебной дисциплины "Математические методы и модели принятия маркетинговых решений" для специальности 1.-28 01 02 "Электронный маркетинг" в БГУИР). Упомянутые и другие отечественные исследования все же лишь косвенно были направлены на проблему автоматизации принятия маркетинговых решений посредством компьютерных программ.

В последние десятилетия был создан большой пласт литературы по автоматизации маркетинга (Marketing automation), которая, однако, больший акцент делает на внедрении систем класса CRM, автоматизации внутрифирменных маркетинговых процессов (в том числе документооборота), а не использовании математико-статистических моделей.

С точки зрения автора, выход в 2017 году книги Ильи Кацова «Введение в алгоритмический маркетинг: искусственный интеллект для маркетинговых операций» [9], если не по содержанию, то по духу, можно считать особой вехой в рассмотренной истории научных исследований. В книге постулируется необходимость синтеза достижений различных наук для создания информационных систем, способных принимать миллионы «микро-решений» (micro-decisions) в маркетинговой деятельности в реальном режиме времени за считанные доли секунды.

Подходя объективно, стоит оценить подход И. Кацова всего лишь как обобщение и систематизацию существующих разработок в теории и практике маркетинга (например, рекомендательные системы, программатик-реклама, динамическое ценообразование и т.п.), однако именно такой синтез позволяет четко зафиксировать новое качество маркетинговой науки – появление алгоритмического маркетинга как реальной прикладной области профессиональной деятельности.

II. НЕКОТОРЫЕ ЗАДАЧИ АЛГОРИТМИЧЕСКОГО МАРКЕТИНГА

В самом широком смысле алгоритмический маркетинг можно определить, как **маркетинговую деятельность, осуществляемую автоматически, без участия (или при самом минимальном участии) человека** [13].

В такой ситуации информационные системы (компьютерное оборудование и программное обеспечение) используют данные о поведении покупателей (в первую очередь об онлайн-поведении), самостоятельно принимают маркетинговые решения, реализуют их и оценивают последствия для достижения целей, сформулированных владельцами бизнеса, топ-менеджерам и/или специалистами по маркетингу.

И. Кацов полагает [9], что в настоящее время алгоритмический маркетинг применим на уровне маркетинговых *процессов*, а не на уровне формулировании маркетинговой *стратегии*. При этом он выделяет следующие 6 областей, в которых достигнут значимый прогресс, – промо-акции (promotions), реклама (advertisements), поиск (search), рекомендации (recommendations), ценообразование (pricing), ассортимент (assortment).

Анализ других научных публикаций позволяет выделить более широкий круг основных областей (задач) алгоритмического маркетинга, в которых в настоящее время ведутся интенсивные научные исследования и прикладные разработки, а также могут ожидать определенные "прорывы". Остановимся на некоторых из них.

Модели реакции рынка (Market Response Models). Классический подход, берущий свое начало еще в середине прошлого

века и подразумевающий моделирование зависимости величины спроса от различных факторов, в первую очередь контролируемых организацией (цены, рекламных усилий и т.п.) с конечной целью оптимизировать расходы на маркетинг для достижения максимальной отдачи.

Полагаем, что в самом общем виде можно формализовать решаемую задачу, как задачу максимизации:

$$R(I) \rightarrow \max_I \quad (1)$$

или как задачу минимизации отклонения получаемых реакций рынка от целевых значений, определяемых собственниками бизнеса и/или топ-менеджментом:

$$|R(I) - G| \rightarrow \min_I \quad (2)$$

Здесь I – вектор «входов» (inputs), т.е. маркетинговых воздействий на рынок (характеристики товара, ассортимент, реклама, цена, упаковка, размещение торговых точек, уровень сервиса, мерчендайзинг и т.п.);

$R(I)$ – функция реакции рынка, которая связывает «входы» с «выходами» (outputs), т.е. измеримыми действиями потребителей (объем и частота покупок, уровень осведомленности о бренде, количество «лайков» в социальных сетях и т.п.). В качестве этой функции в теории широко применяются степенная, экспоненциальная, логистическая функции, а также иные специальные функции, например, ADBUDG-функция Дж. Литтла (John D. C. Little);

G – вектор устанавливаемых собственником бизнеса или топ-менеджерами маркетинговых целей (объем и частота покупок, уровень осведомленности целевой аудитории о товаре, число просмотров веб-страницы, величина прибыли от реализации продукции, доля рынка и т.п.).

Представляется, что модели реакции рынка правильно рассматривать как самую общую задачу алгоритмического маркетинга, которая далее может быть в том или ином виде сформулирована в форме частных задач.

Прогнозирование выбора потребителя (Customer Choice Prediction). Это методы и алгоритмы, позволяющие спрогнозировать предпочтения потребителя относительно ряда (как правило, конкурирующих друг с другом) товаров и услуг, на основе характеристик (явных и неявных) этих товаров и услуг и поведения покупателей в прошлом.

Например, в самом простом случае вероятность того, что n -й покупатель приобретет i -й товар из всех доступных ему на рынке альтернатив (общее число товаров – J), может быть представлена, как:

$$P_{ni} = \frac{e^{A_{ni}}}{\sum_{j=1}^J e^{A_{nj}}}, \quad (3)$$

где P_{ni} – искомая вероятность;

A_{nj} – значение функции полезности n -го покупателя по j -му товару.

Функция полезности в свою очередь может определяться (и оцениваться на основе исторических данных), как:

$$A_{nj} = \sum_{s=1}^S b_s X_{njs} + e_{nj}. \quad (4)$$

где b_s – весовой коэффициент s -го свойства (характеристики) представленных к выбору товаров (например, цвет, мощность, дизайн и т.п.);

X_{njs} – оценка n -го покупателя j -го товара по s -му свойству (характеристике);

e_{nj} – случайная компонента.

Динамическое ценообразование (Dynamic Pricing). Это алгоритмы и системы, осуществляющие варьирование ценой товара или сервиса с целью максимизации получаемой выручки или прибыли. С динамическим ценообразованием тесно связано управление доходами (Revenue Management). Последнее, как правило, относят к сфере ресурсов, предложение которых ограничено (авиаперевозки, гостиничные услуги и т.п.). Некоторые авторы не разделяют, а отождествляют названные области.

В качестве иллюстрации рассмотрим общую постановку задачи в случае, когда целью ценовой политики заявлена максимизация прибыли (мы опускаем для простоты индексы t , хотя нужно понимать, что решается не статическая, а динамическая задача):

$$Profit = (p - VC)Q - FC \rightarrow \max_p. \quad (5)$$

где VC – переменные издержки на единицу продукции (в случае интернет-магазина это может быть, например, закупочная цена единицы продукта, стоимость хранения и доставки этого конкретного продукта);

FC – постоянные издержки на весь объем продаж (в случае интернет-магазина это может быть, например, стоимость аренды офиса, зарплата диспетчера и кладовщика, затраты на хостинг и т.п.).

Тогда для самого простого случая аппроксимации спроса со стороны покупателей посредством линейной регрессии максимизирующее прибыль значение цены будет равно:

$$P^* = \frac{a + bVC}{2b}. \quad (6)$$

где a , b – коэффициенты соответствующей линейной функции спроса.

Прогнозирование оттока клиентов (Churn Prediction). Это алгоритмы, позволяющие на основе данных о покупателях и их прошлых транзакциях (покупках) предсказать вероятность их ухода (отмены подписки, уход к конкуренту и т.п.) в ближайшем будущем и, следовательно, противодействовать этому посредством разнообразных маркетинговых приемов (скидки, распродажи, реклама и т.п.). В самом общем виде можно сформулировать эту задачу как скоринговую модель:

$$S_i = Score(x_i) = Pr(churn|N, x) \quad (7)$$

$$S_i = Score(x_i) = Pr(churn|T, x) \quad (8)$$

где $Pr(churn|N, x)$, $Pr(churn|T, x)$ – вероятность ухода клиента в случае отсутствия (N) и наличия (T) действий со стороны компании по его удержанию (“treatment”) и при заданном векторе характеристик и действий потребителя x .

Рекомендательные системы, системы рекомендаций (Recommender Systems). Это информационные системы (как правило, в форме функционального модуля интернет-магазинов), формирующие список товаров, которые могут быть интересны потребителю, исходя из его прошлых покупок и/или профиля. В настоящее время объектами таких рекомендаций выступают фильмы, музыка, книги, новости, веб-сайты и т.п.

Рекомендации базируются либо на схожести товаров между собой (фильтрация на основе содержания, content-based filtering), либо на схожести потребителей (коллаборативная фильтрация, collaborative filtering), либо используется смешанный подход (hybrid recommender systems).

В качестве иллюстрации расчета подобия a -го и b -го потребителей ($w_{a,b}$) приведем классическую формулу, используемую в коллаборативной фильтрации, базирующуюся на расчете коэффициента корреляции Пирсона:

$$w_{a,b} = \frac{\sum_{i=1}^m (r_{a,i} - \bar{r}_a) \cdot (r_{b,i} - \bar{r}_b)}{\sigma_a \cdot \sigma_b} \quad (9)$$

где m – количество рассматриваемых (и оцениваемых) товаров;

$r_{a,i}$ – оценка a -го потребителя i -го товара;

$r_{b,i}$ – оценка b -го потребителя i -го товара;

\bar{r}_a – средняя оценка a -го потребителя по совокупности всех m товаров;

\bar{r}_b – средняя оценка b -го потребителя по совокупности всех m товаров;

σ_a, σ_b – среднее квадратическое отклонение оценок a -го и b -го потребителей соответственно.

Тогда предсказание оценки a -го потребителя для нового для него j -го товара будет выра-

жаться следующим образом:

$$p_{a,j} = \bar{r}_a + \frac{\sum_{b=1}^n (r_{b,j} - \bar{r}_b) \cdot w_{a,b}}{\sum_{b=1}^n w_{a,b}} \quad (10)$$

где n – количество потребителей, рассматриваемых как подобные для a -го потребителя.

Автоматическая закупка рекламы (Programmatic Advertising). Это алгоритмы и системы автоматической покупки и размещения интернет-рекламы, направленные на максимизацию ее эффективности. Важным направлением programmatic advertising является так называемый торг в реальном времени (Real-Time Bidding), когда автоматическим образом осуществляется купля-продажа рекламных объявлений в реальном времени на аукционе.

Анализ потребительской корзины (Market Basket Analysis). Данная задача предполагает анализ покупок посетителей офлайн- и онлайн-магазинов на предмет выявления устойчивых сочетаний товаров, покупаемых одновременно. Задача, как представляется, носит дополняющий характер, так как ограничивается фиксацией исторических фактов, однако может служить "входом" для решения иных маркетинговых задач.

Прогнозирование покупки (Purchase Prediction). Это алгоритмы, позволяющие по характеристикам и первоначальным действиям посетителя интернет-магазина (или иного вида торговой точки) предсказать, совершит ли он в текущий визит покупку или нет [14], а, следовательно, позволяющих автоматически выбрать наиболее подходящий маркетинговый стимул (персональная скидка, рекомендуемый товар и т.п.) для этого посетителя. Интересен в данном контексте патент компании *Amazon*, в рамках которого предполагается покупка в том или ином географическом регионе и на основании этого прогноза (а не фактического заказа) осуществляется доставка товара с уточнением конечного адреса уже в процессе самой доставки [15].

В качестве иных задач, решаемых в рамках алгоритмического маркетинга, можно также назвать планирование ассортимента (Assortment Planning), прогнозирование популярности контента (Content Popularity Prediction). Особняком стоят задачи, которые скорее должны быть отнесены к области стратегического маркетинга, но тем не менее имеющие потенциал трансформации в технологии и системы алгоритмического маркетинга: сегментирование и кластеризация целевых потребителей, отбор целевых потребителей, моделирование характеристик нового продукта и предсказание его успешности на рынке, оптимизация портфеля товаров и т.д.

III. ЗАКЛЮЧЕНИЕ

Очевидно, что аппарат алгоритмического маркетинга еще будет развиваться, а ИТ-компании будут предлагать готовые технические решения (информационные системы и модули), позволяющие реализовать автоматическое принятие маркетинговых решений в практике существующих компаний (в первую очередь, онлайн-магазинов). Важной задаче выступает включение изучения основ алгоритмического маркетинга в учебный процесс подготовки не только технических специалистов ИТ-профиля, но и менеджеров, маркетологов, экономистов.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Kotler, P. Marketing decision making: a model building approach. – Holt, Rinehart and Winston, 1971 – Business and Economics – 720 p.
2. Lilien, G. L., Kotler, P. Marketing decision making: a model-building approach. – Harper and Row, 1983 – 875 p.
3. Lilien, G. L., Kotler, P., Moorthy, K. S. Marketing Models. Prentice-Hall, 1992 – 803 p.
4. Lilien, G. L., Rangaswamy, A. Marketing Engineering: Computer-assisted Marketing Analysis and Planning. – DecisionPro, 2004 – 518 p.
5. Lilien, G. L., Rangaswamy, A., De Bruyn, A. Principles of Marketing Engineering and Analytics. – DecisionPro, Apr 17, 2017 – 328 p.
6. Wierenga, B. Handbook of Marketing Decision Models. – Springer Science and Business Media, Sep 5, 2008 – 630 p.
7. Hanssens, D. M., Parsons, L., Schultz, R. L. Market Response Models: Econometric and Time Series Analysis. – Springer Science and Business Media, Apr 11, 2006 – 502 p.
8. Wierenga, B., Van der Lans, R. Handbook of Marketing Decision Models. – Springer, Jul 12, 2017 – 598 p.
9. Katsov, I. Introduction to Algorithmic Marketing: Artificial Intelligence for Marketing Operations. – Dec 2, 2017 – 506 p.
10. Малкольм, М. Стратегическое планирование маркетинга. – СПб: Питер, 2000. – 320 с.
11. Анохина, Н. Н. Математические модели маркетинга : Учеб. пособие для вузов / Н.Н.Анохина. – Мн. : БГ-ЭУ, 2002. – 63 с.
12. Сак, А. В., Журавлев, В. А. Оптимизация маркетинговых решений. – Мн.: Издательство Гревцова, 2010. – 304 с.
13. Пархименко, В. А., Шилин, Л. Ю. Концепция системы автоматического управления маркетингом в интернете // Международная научно-техническая конференция, приуроченная к 50-летию МРТИ-БГУИР (Минск, 18-19 марта 2014 года) : материалы конф. В 2 ч. Ч. 2. – Минск, 2014. – С. 247-249.
14. Parkhimenka, U., Tatur, M., Zhvakina, A. Heuristic approach to online purchase prediction based on internet store visitors classification using data mining methods // Information and Digital Technologies (IDT): International Conference. Zilina, Slovakia, July 5-7, 2017.
15. Method and system for anticipatory package shipping [Electronic resource] : pat. US 8615473B2 / Joel R. SpiegelMichael T. McKennaGirish S. LakshmanPaul G. Nordstrom. – Publ. date 17.12.2004. – Mode of access: <https://patents.google.com/patent/US8615473B2/en>.

Секция «Системы управления»

Section «Control systems»

СИСТЕМА АНАЛИЗА И ОЦЕНКИ РИСКОВ СТРАХОВАНИЯ

Акинфина М. А., Бондаренко С. П.

Кафедра информационных технологий, Белорусский государственный экономический университет

Кафедра технологий программирования, Белорусский государственный университет

Минск, Республика Беларусь

E-mail: {akinфина}@rambler.ru, {svetabon}@gmail.com

В данной статье предлагается проект информационной системы для анализа и оценки рисков страхования и реализация средств защиты данной системы.

ВВЕДЕНИЕ

Сегодня мы являемся свидетелями активного развития технологии интеллектуального анализа данных (ИАД или data mining), появление которой связано, в первую очередь, с необходимостью аналитической обработки сверхбольших объемов информации, накапливаемой в современных хранилищах данных.

С помощью ИАД страховщику удастся лучше координировать работу агентов и продавцов и за счет этого повышать эффективность актуарных и андеррайтинговых операций [1].

Целью работы является – проектирование и предложение решения задачи для анализа и оценки рисков страхования, используя самые современные технологии и средства, предлагаемые для реализации информационных систем, в том числе и технологию разведки данных.

I. ПРОЕКТ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Приведем проект системы поддержки и принятия решений оценки и анализа рисков страхования [2].

Система имеет многоуровневую архитектуру и состоит из следующих компонент:

Insurance Database представляет собой реляционную СУБД, в которой реализуется модель данных системы. Вся информация о страховой деятельности фирмы хранится в этой базе, основная задача которой – служить источником данных для корпоративного хранилища;

Integration Service предназначен для миграции данных из базы компании в хранилище, а также отвечает за преобразование информации к требуемому формату, выполняя функции по ее очистке и корректировке;

DataWarehouse – основное хранилище данных компании. Служит для накопления информации о деятельности фирмы за весь период существования. Проектируется с расчетом на необходимость обеспечения удобного анализа хранимой информации. Является ключевым компонентом системы;

Analysis Service служит для построения многомерных моделей на основании данных из хранилища. Предоставляет средства для интел-

лектуального анализа данных, таких как OLAP, Data Mining;

Web Server включает в себя сервис для составления отчетов и корпоративный сайт компании. Обеспечивает доступ к сервисам и данным системы для конечных пользователей.

Для быстрого и разностороннего анализа больших объемов данных спроектировали хранилище данных. Для построения хранилища данных и аналитических сервисов в системе используется Microsoft SQL Server 2005.

Актуарные расчеты – процесс, в ходе которого определяются расходы, необходимые для страхования. С помощью актуарных расчетов определяется стоимость страховой услуги. Как в любой хозяйственной деятельности, в страховании страховщик нуждается в определении размера расходов, необходимых на страхование того или иного объекта. Форма, в которой представляются расходы на страхование данного объекта, называется страховой (актуарной) калькуляцией.

В актуарных расчетах широко используется страховая статистика, которая представляет собой систематизированное изучение и обобщение наиболее массовых и типичных явлений в страховании и их изменение во времени.

Для определения расчетных показателей страховой статистики используются следующие исходные данные:

- число объектов страхования;
- число страховых событий;
- число пострадавших объектов в результате страховых событий;
- сумма собранных страховых платежей;
- сумма выплаченного страхового возмещения;
- страховая сумма для любого объекта страхования;
- страховая сумма, приходящаяся на поврежденный объект страховой совокупности.

После настройки источника данных и конфигурации метаданных получаем OLAP – куб со следующими измерениями:

- Customer – клиенты страховых компаний;
- Risk – описывает возможный страхуемый риск;

Product – представляет страховой продукт компании, предлагаемый клиентам;

Police – заключенный договор между страхователем и компанией;

Claim – описывает наступление страхового случая, Payment – средства полученные от клиентов как страховые премии;

ClaimPayment – средства выплачиваемые компанией как страховые компенсации клиентам;

Location – описывает географическое положение;

Time – время заключения договора.

Приведенная модель позволяет описать всевозможные риски подлежащие страхованию. На основе рисков составляются страховые продукты предлагаемые потенциальным клиентам. Каждый продукт покрывает определенное количество страховых рисков, а также имеет предварительно рассчитанный базовый тариф на данный вид услуги.

При заключении договора за клиентом оформляется страховой полис а также определяются размеры и виды страховых взносов. Также в системе хранится история наступления страховых случаев и проведенных по ним выплат.

Средства Data Mining позволяют аналитикам страховой компании еще более глубоко подойти к оценке стоимости страховых услуг.

Так на основе прошлого опыта аналитик может получить разбиение всех клиентов на схожие группы и оценить выгодность для компании каждой их них. В итоге можно снизить накладные расходы путем повышения ставок для рискованных групп, а также снизить тарифы для благоприятных клиентов с целью повышения спроса на страховые услуги.

В системе все пользователи подразделяются на следующие роли:

–страховой агент, имеющий ограниченный доступ к данным, может использовать инструменты для моделирования и анализа страховых рисков;

–аналитик, который имеет полный доступ к данным для анализа;

–руководитель, имеющий возможность изменения базы данных;

–администратор, который имеет возможность добавление/удаление пользователей, администрирование прав доступа. .

Управление доступом на основе ролей представляет собой модель авторизации, в центре ко-

торой находится пользователь. Вместо перечисления объектов в системе для каждого пользователя и назначения полномочий такое управление доступом позволяет администраторам реализовать управление с точки зрения организационной структуры компании.

Управление доступом на основе ролей обеспечивает центральный объект - роль, назначаемую пользователю для выполнения конкретной функции. Роль явно подразумевает набор разрешений для доступа к некоему набору ресурсов.

При управлении доступом на основе ролей разрешения выдаются не посредством низкоуровневых прав, а высокоуровневых абстракций, соответствующих операциям и заданиям приложения. Операции являются неделимыми единицами, тогда как задания могут состоять из нескольких операций (и других заданий).

II. ЗАКЛЮЧЕНИЕ

В данной работе были изучены основы проектирования и реализации систем поддержки принятия решений. Проанализированы и сопоставлены существующие статистические и эконометрические методики оценки страховых рисков, выявлены лучшие модели.

Внедрение данной системы позволит более точно проработать структуру бизнеса, безошибочно оценить реальный уровень рисков и соответствующим образом назначать цены этих рисков.

Позволит объединить данные в центральном репозитории управления информацией и затем использовать сложные интегрированные приложения для рассмотрения и анализа данных по-новому.

Также позволит обеспечить гибкость в представлении отчетности компании и улучшит представление механизма мониторинга бизнес-процессов, их оперативного и стратегического анализа и планирования.

1. Информационно-аналитические системы в страховании [Электронный ресурс]. – Режим доступа: <http://www.topsbi.ru/default.asp?artID=1126>. – Дата доступа: 25.04.2018.
2. Акинфина, М.А. Использование технологии Data Mining для оценки и контроля рисков кредитования / М.А. Акинфина, С.П. Бондаренко. // Экономический рост Республики Беларусь: глобализация, инновационность, устойчивость: Материалы III Международной научно-практической конференции, Минск, 19-20 мая 2010 г.: В 2 ч. / БГЭУ. — Минск. — 2010. —Ч.2. — С.271–271.

УПРАВЛЕНИЕ ЗАПРОСАМИ К БАЗЕ ДАННЫХ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНОЙ СЕТИ ЧЕРЕЗ ПАТТЕРН MVC

Акиншева И. В., Пантюхов В. А.

Кафедра автоматизации технологических процессов и производств, Могилевский государственный университет продовольствия
Могилев, Республика Беларусь
E-mail: starrina@mail.ru

Разработанное сетевое приложение (социальная сеть) было спроектировано на архитектурном паттерне MVC (Model, View, Controller), который позволил расширить проект от локального узла до многопользовательской сетевой системы. Определен сценарий (механизм выработки) запросов. Для управления запросами использован язык PHP.

ВВЕДЕНИЕ

Используемый паттерн MVC, представляет собой схему разделения данных приложения, пользовательского интерфейса и управляющей логики на три отдельных компонента: модель, представление и контроллер. Модификация каждого компонента осуществляется независимо [1].

Входной точкой в разработанное программное обеспечение, реализующее социальную сеть, является контроллер, который реагирует на действие пользователя и отправляет соответствующий ответ. Причем между этими двумя действиями контроллер может обратиться как к модели, так и к сервисам, которые делегируют выполнение работы по обработке данных моделям. Пользовательский интерфейс с точки зрения MVC называется View (Представление). Входная точка в приложение возвращает View. То, что каждый запрос возвращает View с инкасулированными в него данными, является не обязательным, это зависит от типа приложения.

Если это SPA (Single Page Application) приложение, то сервер возвращает View один раз, а все остальные запросы просто отправляют нужные данные клиентской части, которые самостоятельно генерирует View внутри себя (чаще всего с помощью фреймворка). Если же это SSR-SPA приложение (Server Side Rendering + Single Page Application), то сервер возвращает View больше одного раза, однако скрипты, формирующие запросы приложению, загружаются не сразу. Разработанное приложение имеет тип SPA. В качестве шаблона для построения серверной части выступает фреймворк Laravel.

I. ОСНОВНАЯ ЧАСТЬ

Используемая в приложении гипертекстовая разметка является минимальной с точки зрения формирования базовой html-структуры.

Запросы происходят по следующему сценарию:

- пользователь входит в приложение, основной контроллер принимает запрос от при-

ложения «клиента», расположенного на машине пользователя;

- контроллер возвращает приложению «клиента» файл с гипертекстовой разметкой (view) и скрипты;
- браузер отображает файл с гипертекстовой разметкой и загружает скрипты;
- при загрузке скриптов в приложение «клиента» оформляется остальная часть пользовательского интерфейса и добавляются элементы интерактивности на страницу в браузере.

Выполнение такой последовательности действий, позволяет пользователю увидеть свой аккаунт в социальной сети как можно быстрее.

Из описанного выше следует, что используемый паттерн MVC является универсальным, не зависит от подхода к разработке приложения и позволяет оперативно изменять конфигурацию приложения.

В данном проекте паттерн MVC реализован полностью, все три сущности изолированы друг от друга благодаря использованию современного фреймворка Laravel, написанного на языке PHP.

Рассмотрим, как реализован паттерн MVC в данном приложении (рис. 1).

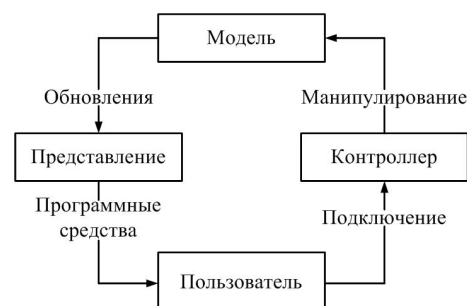


Рис. 1 – Диаграмма MVC, используемая для формирования приложения

В первую очередь запрос попадет в файл routes/web.php, который является точкой входа для любого запроса в данном приложении. Файл содержит имена всех роутеров и контроллеров,

которые отвечают за используемые и выполняемые запросы.

В разработанном приложении посылается запрос по корневому адресу приложения, за который отвечает представленный ниже элемент файла web.php:

```
Route::get('/', function () { // Главная страница
    return view('welcome'); // показать файл welcome.blade.php
})->name('home'); // Имя роутера home
```

В данном случае не описывается, какой конкретно контроллер отвечает за логику обработки запроса, поскольку запрос обрабатывается прямо внутри представленного блока кода, возвращая «view» с именем «welcome». Интерпретатор воспринимает, что полное название файла будет welcome.blade.php [2].

Созданный файл welcome.blade.php является так называемым «view engine». Это обычный файл с расширением php, который интерпретатор превращает на выходе в файл с форматом .html, предварительно вставив в него динамическую информацию и сгенерировав динамическую html-структуру [3]. Например, для пользователя, написавшего два сообщения будет сгенерирована структура, содержащая два html-блока с различным содержанием, но одинаковой html-структурой. После формирования окончательной html-структуры, интерпретатор возвращает файл с расширением .html и сервер возвращает данный файл приложению «клиента».

Рассмотрим для примера один из контроллеров, реализованных в разработанном приложении. Сам файл представляет собой класс, наследующий класс фреймворка Laravel, который реализует различные интерфейсы. Класс UserController инкапсулирует в себе публичные методы, задающие логику обработки запросов, привязанных конкретно к используемому контроллеру. Этот механизм реализован в файле web.php. Пусть путь к post-запросу прописывается следующим образом: «socialnet.loc/signin». Этот post-запрос предназначен для авторизации пользователя. Запрос поступает в файл web.php, где обрабатывается следующим блоком кода:

```
Route::post('/signin', [
    'uses' => 'UserController@postSignIn',
    'as' => 'signIn' ]);
```

Во второй строке кода реализована логика, задающая контроллер конкретно для этого запроса. Также делегируется возможность использования метода ответственным за обработку запроса типа post. Контроллер с именем UserController выполняет метод postSignIn. «As» используется для сокращенного обращения к классу UserController через view. Используя сокращение, появляется возможность активировать этот контроллер через любое view следующим образом:

```
<form action="{{ route('signIn') }}" method="post" >html код </form>
```

Данная строка обрабатывается компилятором и на выходе получается следующий код:

```
<form action="/signin" method="post" >html код </form>
```

В общем случае считается хорошей практикой использовать такого рода сокращения методов. Запросы могут быть длинными, а код желательно изначально делать однотипным в пределах всего приложения. Кроме того в любом месте кода можно активировать данный запрос в выражении «route('signIn')».

При создании модели данных, в представленном случае это будет пользователь, необходимо создать для нее класс и установить связи с другими моделями данных. В данном приложении используются связи «один ко многим», «многие ко многим», «один к одному». Рассмотрим пример, когда пользователь имеет несколько постов. На программном уровне это инициализируется следующим образом:

```
class User extends Model implements Authenticatable {
    public function posts() {
        return $this->hasMany('AppPost');
        //Relations!! }}

```

Реализация связи «многие к одному» приводит к тому, что пользователь может иметь несколько постов, в свою очередь пост имеет одного пользователя:

```
class Post extends Model {
    public function user() {
        return $this->belongsTo('AppUser'); }}

```

Согласно приведенному коду, при компиляции автоматически создаются таблицы в базе данных с требуемыми связями. Таким образом, разработчику не нужно знать все нюансы синтаксиса SQL для создания связей между отдельными таблицами в базе данных.

II. ЗАКЛЮЧЕНИЕ

В результате было разработано сетевое приложение на архитектурном паттерне MVC со следующими возможностями: валидация данных на стороне «сервера»; валидация данных на стороне «клиента»; применение современных UI-стандартов со стороны «клиента»; защищенное соединение с административной панелью, расширенные возможности работы с интерфейсом страницы пользователя.

1. Стивенс, У. Р. UNIX: разработка сетевых приложений / У. Р. Стивенс, Б. Феннер, Э. М. Рудолфф. – СПб.: Питер, 2007. –1039 с.
2. Котеров, Д. PHP 7 / Д. Котеров, И. Симдянов. – СПб.: БХВ-Петербург, 2016. –1088 с.
3. Бибо, Б. jQuery. Подробное руководство по JavaScript / Б. Бибо, И. Кац; пер. с англ. – СПб.: Символ-Плюс, 2011. –624 с.

СИСТЕМА ОТОБРАЖЕНИЯ ИНФОРМАЦИИ НА БАЗЕ КОНТРОЛЛЕРА STM И ПЛАТФОРМЫ ANDROID

Бранцевич П. Ю., Деменковец Д. В.

Кафедра программного обеспечения информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: demenkovets@bsuir.by

В докладе рассмотрены структурная схема системы, устройство отображения, описание функций программного средства конфигурирования устройства отображения.

ВВЕДЕНИЕ

Эпоха статической информации уходит, все большее внимание привлекают информационные табло с динамическим отображением информации.

Рекламные щиты и плакаты уступают место мультимедийной и динамичной рекламе. Мультимедийная наружная реклама подчеркивает индивидуальность отдельной компании, выделяет ее на фоне других. Появляется возможность оперативного изменения рекламной информации.

Сегодня монохромная светодиодная бегущая строка или информационное табло – простейший из ряда инструментов динамической рекламы. Данный вид рекламных технологий позволяет выводить на экран информацию из цифр и букв.

В связи с этим была поставлена цель разработать систему отображения информации с разработкой программного обеспечения устройства отображения и программного средства для его конфигурации. В данном докладе рассмотрена организация разработанной системы.

I. ОРГАНИЗАЦИЯ СИСТЕМЫ ОТОБРАЖЕНИЯ ИНФОРМАЦИИ

Была разработана система отображения информации состоящая из следующих компонентов:

- устройства отображения;
- мобильного устройства конфигурирования под управлением операционной системы Android.

В свою очередь устройство отображения состоит из следующих компонентов:

- микроконтроллер STM32F103;
- датчик температуры и влажности HTU-21D;
- дисплей на базе контроллера MAX7219;
- модуль Bluetooth LE HM-10;
- блок питания.

Структурная схема системы отображения изображена на рисунке 1.

Взаимодействие мобильного устройства конфигурирования с устройством отображения

выполняется посредством беспроводной связи по технологии Bluetooth LE [1].

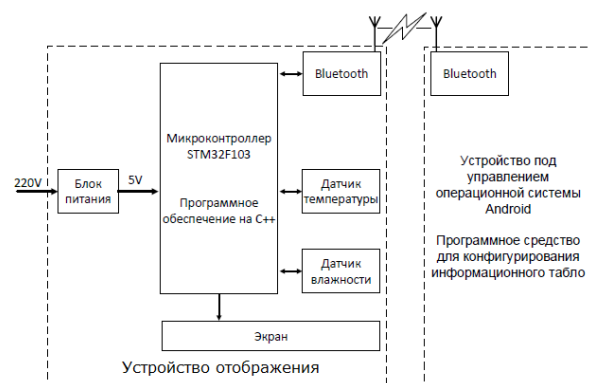


Рис. 1 – Структурная схема системы отображения информации

Основные проблемные вопросы возникающие при создании подобных систем являются:

- поддержка русского языка;
- управление динамикой отображения информации;
- организация доступа к устройству отображения в соответствии с правами;
- корректировка показаний с датчиков влажности и температуры;
- степень защиты от окружающей среды и антивандальность;
- возможность конфигурации с мобильного устройства (смартфона) без использования проводного соединения;
- пользовательский интерфейс на русском языке;
- энергопотребление;
- стоимость.

Основные проблемные вопросы удалось решить в разработанном программном обеспечении под платформу Android, с разработанным протоколом взаимодействия между устройством отображения и устройством конфигурирования.

В предложенной системе поддержка русского языка решена с помощью алгоритма формирования символа изображенного на рисунке 2.

Корректировка показаний температуры и влажности осуществляется с помощью записи поправочных коэффициентов в энергонезависимую память контроллера.

Степень защиты устройства отображения предложено решить с помощью корпуса, который выполнен из металла.

Организация прав доступа к устройству осуществляется с помощью паролей и уровней доступа.

Для уменьшения энергопотребления предложен оптимальный алгоритм работы контроллера и выбран модуль bluetooth с пониженным энергопотреблением.

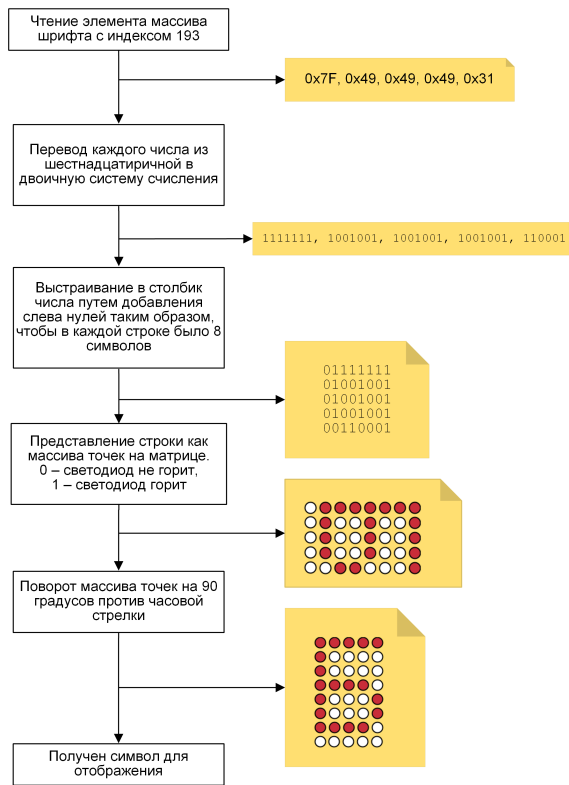


Рис. 2 – Алгоритм формирования символа для отображения

II. ПРОГРАММНОЕ СРЕДСТВО СИСТЕМЫ И УСТРОЙСТВО ОТОБРАЖЕНИЯ

Программное средство системы позволяет производить конфигурацию и управление устройством отображения в соответствии с правами доступа.

Технические характеристики и возможности ПС и системы:

- поддержка уровней доступа к устройству отображения с мобильного устройства (гость, пользователь, администратор);
- авторизация с помощью пароля;
- управление с мобильного устройства;
- конфигурирование устройства с помощью беспроводного интерфейса (bluetooth);
- сохранение конфигурации в энергонезависимой памяти;
- отображение текущих показаний датчиков температуры и влажности;

- возможность изменения скорости движения информационного сообщения в режиме «бегущая строка»;
- управление пользовательским сообщением;
- степень защиты устройства отображения IP65;
- антивандальный защитный корпус.

Разработка программы контроллера STM32F103 выполнена на языке C++ с среде разработки Keil uVision[2]. В качестве языка программирования для программного средства по конфигурированию информационного табло был выбран язык Kotlin. Данный язык был представлен как официальный язык разработки для платформы Android 17 мая 2017 года [3,4]

III. ЗАКЛЮЧЕНИЕ

Данная система может иметь широкое применение. Внешний вид устройства отображения и вариант его использования представлен на рисунке 3.



Рис. 3 – Внешний вид устройства. Вариант применения

IV. СПИСОК ЛИТЕРАТУРЫ

1. Bluetooth Low Energy [Электронный ресурс] / – Режим доступа: <http://www.mt-system.ru/catalog/bluetooth-low-energy/> – Дата доступа: 16.09.2018.
2. Arm Keil Vision® IDE [Электронный ресурс] / – Режим доступа: <http://www2.keil.com/mdk5/uvision/> – Дата доступа: 16.09.2018.
3. Android Announces Support for Kotlin [Электронный ресурс] / – Режим доступа: <https://android-developers.googleblog.com/2017/05/android-announces-support-for-kotlin.html> – Дата доступа: 16.09.2018.
4. Kotlin Documentation [Электронный ресурс] / – Режим доступа: <https://kotlinlang.org/docs/reference/> – Дата доступа: 16.09.2018.

МЕТОД ЛОГИКО-ВРЕМЕННОГО АНАЛИЗА ЦИФРОВЫХ СИСТЕМ С ОГРАНИЧЕНИЕМ НА ВРЕМЯ ОТКЛИКА

Черемисинов Д. И

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: cher@newman.bas-net.by

Предлагается метод анализа цифровых систем по алгоритмическому описанию на языке параллельных алгоритмов логического управления. Метод основан на построении множества структурных состояний переходов параллельного алгоритма и расширенного графа достижимости состояний на множестве меток и логических переменных алгоритма, и является модификацией метода символической проверки на модели.

ВВЕДЕНИЕ

Интегральные схемы в настоящее время проектируются в результате выполнения ряда шагов, каждый из которых переводит абстрактную спецификацию, уточняя ее, во все более конкретное воплощение. Процесс проектирования начинается с «поведенческой модели», например, с программы, которая описывает архитектуру процессора на уровне набора команд. Результатом проектирования является описание фактической топологии транзисторов и проводов на чипе. Каждый шаг проектирования должен давать описание, релевантное исходному поведению абстрактной модели. К сожалению, методы выполнения проверки соответствия спецификации и реализации становятся все более дорогими и трудными по мере роста сложности проекта. Для многих проектов размер группы проектировщиков, выполняющих проверку (верификацию) проекта, теперь превышает число участников группы проектирования.

Формальная проверка обеспечивает новый подход к подтверждению правильного поведения описания СБИС. В моделировании, традиционном способе отладки описания СБИС, подтверждение правильности текущего описания СБИС – это результат прогонов моделирования с большим количеством испытательных ситуаций. Формальная проверка (верификация), напротив, использует математические методы, чтобы исследовать полное пространство состояний модели для подтверждения соответствия наперед заданному поведению. последние годы инструменты для символической верификации находят существенное и растущее применение, они рутинно используются в промышленных САПР [1].

Ключевой вопрос при использовании инструмента для символической верификации – это формулировка свойств - assertion на языке спецификации временных условий, который является одним из определяющих интерфейсов программы. (Другой важный интерфейс – это язык моделирования, в этом качестве обычно используется язык описания аппаратных средств

ЭВМ, применяемый проектировщиками СБИС). Языки спецификации временных условий имеют в основе формализм временных логик [2, 3].

Алгоритмы верификации на символической модели могут проверять большой и важный класс свойств цифровых систем. Тем не менее, существует важный класс свойств, который нельзя адекватно обрабатывать с помощью этого метода. Этот класс состоит из свойств, которые ограничивают время отклика. Во временных логике CTL [3] можно указать, что какое-то событие произойдет в будущем, но свойство, что какое-то событие произойдет не более чем через x единиц временных не может быть выражено напрямую.

В настоящей работе метод символической верификации, модель проверяемого устройства задается на языке ПРАЛУ [4], и метод дает минимальный и максимальный временной интервал для перехода из известного состояния запуска ПРАЛУ алгоритма в заданное.

I. ПРЕДЛАГАЕМЫЙ МЕТОД

Проверяемая система задается на языке ПРАЛУ и затем компилируется в граф переходов состояний. алгоритм на ПРАЛУ служит точным описанием системы, которое может выявить тонкие двусмысленности и может использоваться для целей документирования. Формализация понятия процесса в языке ПРАЛУ использует в качестве базы представление об операции. Временное свойство всех операции одно: протяженности во времени – операция это интервальное событие, т.е. временной интервал конечной, но не нулевой длины. Предполагается, что все элементарные операции действия имеют одинаковую длительность, т.е. параллельные процессы выполняются в режиме блокировки (lock-step – шагать в ногу тесным строем).

Метод позволяет получить количественную информации о времени отклика, т.е. минимальные и максимальные задержки между запросом и соответствующим ответом. Например, метод позволяет получить границы интервала времени между запросом на доступ к шине и соответству-

ющим предоставлением доступа к шине. Кроме того, можно вычислить количество случаев, когда третье событие происходит в течение такого интервала, например, количество других регистровых пересылок, происходящих между запросом на шину и соответствующим разрешением на доступ.

Основным ограничением нашего подхода является неотъемлемая сложность проблемы проверки на модели. Граф переходов имеет экспоненциальную асимптотическую сложность в количестве вершин, и нет никаких гарантий того, что алгоритм анализа сможет выполнить проверку в любом практическом примере. В большинстве случаев проверка выполняется в считанные минуты, даже для сложных систем реального мира. Следует, однако, сказать, что эти проблемы присущи любому формальному методу логико-временного анализа.

II. СКЕЛЕТ α -СЕТИ

Далее для задания поведения систем управления используется параллельные алгоритмы логического управления на языке ПРАЛУ в стандартном виде модели параллельного автомата [4]. Алгоритмы в таком виде представляют подкласс раскрашенных сетей Петри – расширенные сети свободного выбора. Алгоритм представляет собой совокупность стандартных цепочек вида: $\tau_i = (\mu_i \rightarrow \nu_i) / (k_i^1 \rightarrow k_i^2)$, где метки τ_i и ν_i трактуются как подмножества частичных состояний, а элементарные конъюнкции k_i^1 и k_i^2 – как условие перехода и выходные сигналы, сопровождающие переход. Переход срабатывает, когда текущая маркировка N_t на множестве состояний $M = 1, 2, \dots, m$ включает все состояния из μ_i и переменные принимают значения, обращающие k_i^1 в 1. После срабатывания перехода переменным из k_i^2 присваиваются значения, обращающие k_i^2 в 1, а маркировка N_t заменяется на $(N_t \setminus \mu_i) \cup \nu_i$. Формальная верификация описания алгоритма на языке ПРАЛУ основана на анализе его скелета, α -сети, задаваемого множеством переходов $\tau_i = (\mu_i \rightarrow \nu_i)$ и соответствующего ординарной сети Петри.

Основной структурой, лежащей в основе методов анализа поведенческих свойств алгоритма на языке ПРАЛУ, является его граф достижимости, который строится на основе α -сети алгоритма управления. Вершинам графа достижимости соответствуют все возможные разметки N_t , а дуга, помеченная символами одного или нескольких переходов, соединяет разметки, такие, что сеть переходит от первой разметки ко второй при срабатывании этих переходов. Анализируя α -сеть и построенный на ее основе граф достижимости, можно исследовать корректность алгоритма на языке ПРАЛУ [5], но это исследование будет неполным. Для полноты необходимо учесть в нем логические характеристики: достижимые структурные состояния на множе-

стве его переменных, которые обуславливают и временные характеристики выполнения алгоритма управления. Соответственно дуги графа достижимости необходимо помечать не только переходами, но и парой: входное условие срабатывания перехода и значения изменяемых переменных. Метод построения такого расширенного графа достижимых частичных состояний и соответствующих достижимых состояний логических переменных основан на получении множества структурных состояний переходов параллельного алгоритма – состояний его логических переменных на момент переходов.

III. КОЛИЧЕСТВЕННЫЙ АНАЛИЗ, ВРЕМЯ ОТКЛИКА

Проверка на модели (Model checking) – это метод автоматического анализа, который исследует все возможные состояния анализируемой системы, чтобы проверить, удовлетворяет ли система формально заданному свойству. Свойства записываются в виде формул в пропозициональной временной логике. В предлагаемом методе свойства задаются неявно указанием пары разметок графа достижимости для вычисления количественной временной информации, такой как точные минимальные и максимальные задержки на время между запросом и соответствующим ответом.

Свойства, необходимые для количественного временного анализа, не могут быть выражены в традиционной пропозициональной временной логике. Можно обеспечить задание и проверку ограниченных по времени свойств введением границ в временные операторы временной логики. Расширенная логика называется RTCTL [6]. Предлагаемый метод задания информации для количественного анализа поведения не так универсален как RTCTL, но позволяет упростить алгоритм проверки на модели для типично проверяемых свойств.

1. J. R. Burch, E. M. Clarke, K. L. McMillan, D. L. Dill, and J. Hwang. Symbolic model checking: 1020 states and beyond. Information and Computation, Volume 98, Issue 2, June 1992, Pages 142-170.
2. L. Lamport. Sometimes is sometimes “not never” – on the temporal logic of programs. In Proc. 7th ACM Symp. on Principles of Programming Languages, pages 174–185, January 1980.
3. F. Wang. Timing behavior analysis for real-time systems. Proceedings of Tenth Annual IEEE Symposium on Logic in Computer Science, San Deigo, CA, USA, 1995, pp. 112-122.
4. Черемисинова, Л.Д. Реализация параллельных алгоритмов логического управления – Минск: Ин-т техн.кибернетики НАН Беларуси, 2002. – 246 с.
5. Закревский, А.Д. Параллельные алгоритмы логического управления – Минск: Ин-т техн.кибернетики НАН Беларуси, 1999. – 202 с.
6. E. A. Emerson, A. K. Mok, A. P. Sistla, and J. Srinivasan. Quantitative temporal reasoning. In Lecture Notes in Computer Science, Computer-Aided Verification. Springer-Verlag, 1990.

К АНАЛИЗУ ПАРАЛЛЕЛЬНОГО АЛГОРИТМА УПРАВЛЕНИЯ НА ОСНОВЕ РАСШИРЕННОГО ГРАФА ДОСТИЖИМОСТИ

Черемисинова Л. Д.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: cld@newman.bas-net.by

Для анализа реальных систем логического управления введено понятие расширенного графа достижимости параллельного алгоритма и приведен метод его построения.

ВВЕДЕНИЕ

Для правильности функционирования системы логического управления необходимо обеспечить предсказуемое время отклика на внешние стимулы. Анализ описания функционирования системы связан с установлением факта наличия у нее некоторого набора желательных свойств [1] и требует исследования логико-временных характеристик его выполнения. Основным инструментом, лежащим в основе методов анализа поведенческих свойств алгоритма управления, является его граф достижимости, анализируя который, можно проверить корректность параллельного алгоритма управления. Однако для полноты анализа необходимо учесть также логические характеристики алгоритма: достижимые состояния на множестве его переменных, которые существенно влияют на порядок выполнения переходов и их срабатывание. В работе вводится понятие расширенного графа достижимости параллельного алгоритма, который содержит достижимые полные состояния не только на множестве его меток, но и на множестве его логических переменных.

I. УЧЕТ ЛОГИКО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК

Для задания поведения систем управления далее используется модель параллельного автомата на языке ПРАЛУ [1], которая представляет подкласс раскрашенных сетей Петри – расширенные сети свободного выбора. Алгоритм представляет собой совокупность цепочек вида: $\tau_i = (\mu_i \rightarrow \nu_i) / (k_i^1 \rightarrow k_i^2)$, где метки μ_i и ν_i трактуются как подмножества частичных состояний, а элементарные конъюнкции k_i^1 и k_i^2 – как условие перехода и выходные сигналы, сопровождающие переход. Срабатывание переходов зависит не только от достигнутого состояния на множестве меток переходов, но и от условий, определенных на множестве переменных (аналогично раскрашенным сетям Петри). Наличие условий обеспечивает возбуждение только тех переходов τ_i , возможных для достигнутой маркировки N_t , для которых выполняется не только условие $\mu_i \subseteq N_t$, но и $k_i^1 \wedge W_t \neq 0$, где W_t – состояние на множестве логических сигналов систе-

мы управления, достигнутое к этому к моменту t . В случае корректного алгоритма для любой пары переходов τ_i и τ_j выполняется условие $(i \neq j) \wedge (\mu_i \cap \mu_j \neq \emptyset) \rightarrow \mu_i = \mu_j$ и, если τ_i и τ_j параллельны, то $\nu_i \wedge \nu_j = \emptyset$. После срабатывания перехода переменным из k_i^2 присваиваются значения, обращающие k_i^2 в 1, а $N_{t+1} = (N_t \setminus \mu_i) \cup \nu_i$.

Динамика сети описывается в пространстве достигаемых состояний переходов и состояний на множестве логических сигналов. Для вычисления следующего состояния необходимо определить множество T переходов $\tau_i = (\mu_i \rightarrow \nu_i) / (k_i^1 \rightarrow k_i^2)$, которые срабатывают при полученной маркировке N_t , и изменение состояния W_t на множестве переменных системы управления. Следующие состояния определяются как:

- 1) $N_{t+1} = (N_t \setminus \mu) \cup \nu$, где $\mu = \cup_i \mu_i$, $\nu = \cup_i \nu_i$ – объединения множеств частичных состояний всех переходов $\tau_i \in T$, сработавших параллельно;
- 2) значения переменных, упомянутых в конъюнкциях k_i^2 переходов $\tau_i \in T$, изменяются на значения, обращающие каждую из k_i^2 в 1.

Вершинам графа достижимости соответствуют все возможные разметки N_t , а дуга, помеченная символами одного или нескольких переходов, соединяет разметки, такие, что сеть переходит от первой разметки ко второй при срабатывании этих переходов. Для учета логических характеристик дуги графа достижимости необходимо помечать не только переходами, но и парой: входное условие срабатывания перехода и значения изменяемых переменных, а также значения состояний логических переменных алгоритма в вершинах графа достижимости.

II. ПОСТРОЕНИЕ СТРУКТУРНЫХ СОСТОЯНИЙ ПЕРЕХОДОВ

Метод построения расширенного графа достижимых частичных состояний и соответствующих достижимых состояний логических переменных основан на получении множества структурных состояний переходов параллельного алгоритма – состояний его логических переменных на момент переходов.

Под структурным состоянием параллельного автомата понимается набор значений всех его переменных в некоторый момент времени. Смена структурного состояния при переходе $\tau_i =$

$(\mu_i \rightarrow \nu_i)/(k_i^1 \rightarrow k_i^2)$ состоит в смене значений только тех переменных, которые входят в конъюнкцию k_i^2 , ими являются выходные или внутренние переменные, составляющие множество W переменных, значения которых контролируются алгоритмом управления. Конъюнкция k_i^2 (определяемая множеством ν_i частичных состояний), задавая изменение значений только упомянутых в ней переменных, не является понимаемым в традиционном смысле структурным выходным состоянием автомата. Для его получения необходимо найти значения остальных переменных из W . При этом искомое выходное состояние перехода τ_i , если $|\nu_i| > 1$, делится на несколько неортогональных составляющих, каждая из них определяется конечными состояниями переходов в состояние $s_{ki} \in \nu_i$.

Перед построением структурных состояний переходов параллельного алгоритма: 1) для каждого s_i определяется множество P_i состояний, ему параллельных и определяется множество $W_i^* = W \setminus Z_i$ переменных, контролируемых состоянием s_i [2], где Z_i – множество переменных, изменяемых при переходах из состояний множества P_i (если алгоритм корректен, то $Z_i \cap W_i = \emptyset$); 2) начальное состояние w_n^i каждого перехода $\tau_i = (\mu_i \rightarrow \nu_i)/(k_i^1 \rightarrow k_i^2)$ определяется на множестве переменных из W , контролируемых состояниями из μ_i , при этом w_n^i состоит из $|\mu_i|$ неортогональных составляющих, каждая из которых определяется конечными состояниями переходов τ_j в частичное состояние $s_{ki} \in \mu_i$ и $s_{kj} \in \nu_j$; 3) конечное состояние w_k^i перехода τ_i отличается от w_n^i значениями только тех переменных, которые входят в k_i^2 .

Вычисление начальных состояний переходов на множестве W осуществляется в процессе моделирования поведения алгоритма управления на множестве T его переходов. В качестве начального состояния первого перехода t_1 принимается структурное состояние автомата на множестве W перед началом его функционирования. На каждом шаге выбирается и исключается из T один корректируемый переход t_i , вычисляются его новые начальное w_n^i и конечное w_k^i состояния на множестве текущих конечных состояний его переходов-предшественников. Если w_k^i изменилось, то множество T пополняется всеми переходами-последователями t_i . Процесс вычисления начальных и конечных состояний продолжается до тех пор, пока множество T не опустеет.

III. ПОСТРОЕНИЕ РАСШИРЕННОГО ГРАФА ДОСТИЖИМОСТИ

Расширенный граф достижимости алгоритма строится с учетом состояний всех его переменных. Его вершинами являются подмножества N_1, N_2, \dots, N_t частичных состояний алгоритма управления и структурные состояния на мно-

жестве значений переменных алгоритма, представленные в виде конъюнкций K_1, K_2, \dots, K_t . Из вершины N_g графа достижимости исходит дуга, заходящая в вершину N_h , если в исходном задании алгоритма имеется переход $\tau_i = (\mu_i \rightarrow \nu_i)/(k_i^1 \rightarrow k_i^2)$, начальным и конечным структурными состояниями которого являются w_n^i и w_k^i , такой, что

- 1) $N_h = (N_g \setminus \mu_i) \cup \nu_i$;
- 2) $K_g \wedge w_n^i \neq 0$;
- 3) $(K_h \wedge w_k^i) \neq 0$.

Метод построения графа достижимости состоит в последовательном получении полных состояний N_t и соответствующих им структурных состояний на множестве значений переменных алгоритма, начиная от заданного начального полного состояния в виде множества N_1 , которое является первым и единственным элементом искомого множества полных состояний. В качестве структурного состояния K_1 на множестве значений переменных принимаются значения переменных из W перед началом его функционирования алгоритма.

Далее производится просмотр всех переходов $\tau_i = (\mu_i \rightarrow \nu_i)/(k_i^1 \rightarrow k_i^2)$ параллельного автомата, и, если $\mu_i \subseteq N_1$ и $K_1 \wedge w_n^i \neq 0$, то формируются новые множества N_j и K_j . $N_j = (N_1 \setminus \mu_i) \cup \nu_i$ объявляется очередным достижимым полным состоянием, а в качестве K_j принимается структурное состояние K_1 , в котором значения всех переменных, упомянутых в k_i^2 изменяются. Полученной паре K_1, K_j достижимых состояний соответствует дуга искомого графа, направленная из K_1 в K_j . Дуга помечается номером перехода τ_i . Этот процесс повторяется для каждого из вновь внесенных в граф достижимости множеств N_j . Процесс заканчивается, когда не получается новых множеств такого вида, отличных от уже полученных.

IV. ЗАКЛЮЧЕНИЕ

Расширенный граф достижимости состояний на множестве меток и логических переменных алгоритма управления позволяет проанализировать последовательность изменения состояний системы управления в ответ на внешние воздействия и обеспечивает более точный логико-временной анализ реальных систем логического управления по сравнению с методами формального анализа описаний алгоритмов на языках, подобных сетям Петри.

V. СПИСОК ЛИТЕРАТУРЫ

1. Закревский, А. Д. Параллельные алгоритмы логического управления – Минск: Ин-т техн.кибернетики НАН Беларуси, 1999. – 202 с.
2. Черемисинова, Л. Д. Реализация параллельных алгоритмов логического управления – Минск: Ин-т техн.кибернетики НАН Беларуси, 2002. – 246 с..

СИСТЕМА АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ В АВТОМОБИЛЯХ

Гинько Д. В., Цингель А. А., Курулев А. П.

Кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: E-mail: denisginko@icloud.com, anton_tsg@mail.ru, alexpapakuru@yahoo.com

В работе приводится описание системы автопилота в автомобилях, ее преимущества и недостатки.

ВВЕДЕНИЕ

Автомобильная индустрия претерпевает существенную трансформацию: крупнейшие производители машин совместно с ИТ и телеком разработчиками идут к созданию транспортных средств с возможностью полностью автономного вождения. Тренд уже очевиден – в будущем беспилотный транспорт станет массовым явлением.

I. ОБЩИЕ ПРИНЦИПЫ РАБОТЫ БЕСПИЛОТНОГО АВТОМОБИЛЯ

В мире идет активная разработка ITS нового поколения с большим спектром возможностей, их стандартизацией занимаются такие организации, как ETSI, IEEE, 3GPP и другие. Современные системы ITS решают такие задачи, как контроль допуска, управление и оплата парковками, предоставление информации о движении и оплате парковки, управление грузоперевозками, контроль трафика и т.д.

Одним из основных применений ITS является помощь водителю транспортного средства. За счет кооперативной осведомленности транспортное средство может получить оповещение об опасности, индикатор медленно идущих машин, предупреждение о столкновении на перекрестке, индикатор о приближении мотоцикла и т.д.

Водителю будут доступны оповещения о таких ситуациях, как поломка электрического освещения, неверная дорога, стационарная машина (авария или поломка транспортного средства), проведение дорожных работ, риск столкновения, оповещение о состоянии дорожного движения и оповещение о смене сигнала. Децентрализованные базы данных будут предоставлять информацию об опасных зонах, осадках, сцеплениях на дорогах, видимости, ветре и др.

Следующим шагом станет использование ITS в беспилотных автомобилях. Базовым компонентом беспилотников будут внешние камеры и радарное оборудование, отмечается в отчете НИИР. Но именно обмен информацией между автомобилями по средствам V2V-систем вместе с получением транспортными средствами через V2I-системы информации о ситуации на дорогах и актуальных цифровых карт дорог позволит

обеспечить безопасное и эффективное дорожное движение беспилотников.

II. Типы ITS: V2V и V2I

Первый тип систем - «транспортное средство - транспортное средство» (vehicle-to-vehicle, V2V) - обеспечивают безопасное вождение за счет связи между автомобилями на перекрестках с плохой видимостью. V2V-система может предупреждать водителей об опасности лобового столкновения, бокового столкновения, заднего столкновения, уведомлять о неисправности транспортного средства, предоставлять дорожную и нормативную информацию

Например, две машины, невидимые друг другу на перекрестке или на повороте, через V2V-систему могут обмениваться друг с другом координатами и значениями скоростей для избежания столкновения. Аналогичным образом автомобиль, приближающийся к концу пробки, получит информацию с координатами и скоростями ближайших транспортных средств.

Второй тип систем безопасного движения - «придорожная инфраструктура - транспортное средство» (vehicle-to-infrastructure, V2I) - обеспечивают передачу информации (сигнал и нормативная информация и т.д.) от придорожного оборудования к автомобилю через средства радиосвязи. Например, придорожные сенсоры на перекрестке обнаружат машины, которые собираются пересечь перекресток или повернуть, и передадут информацию другим приближающимся машинам по средствам V2I-систем.

III. Степени автономности автомобилей

По классификации SAE International систем помощи водителю или ADAS (Advanced Driver Assistance System) существует шесть классов автономности от уровня 0 — полностью ручное управление с возможностью предупреждения об опасных ситуациях на дороге, до 5 — полностью беспилотный автомобиль.

Уровень ADAS 1 предусматривает работу более продвинутой системы предупреждения об опасности столкновения с автомобилями, пешеходами, а также о пересечении линии разметки, идентификацию дорожных знаков и т. п., а также вмешательство в систему управления.

Второй уровень — это более активная помощь водителю (руление, торможение, удержание в полосе и т. д.).

Третий уровень — автономное движение на заданных участках дороги, которое требует от водителя лишь частичного надзорного контроля.

Четвертый уровень ADAS — это автономное движение автомобиля в определенных режимах, при котором человек уже не может повлиять на управление даже в критических ситуациях.

Пятый уровень — полная автономность транспортного средства, когда водитель лишь задает ко-нечный пункт маршрута, а весь процесс передвижения полностью ложится на искусственный интеллект автопилота.

IV. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ

Преимущества:

- перевозка грузов в опасных зонах, во время природных и техногенных катастроф или военных действий.
- снижение стоимости транспортировки грузов и людей за счёт экономии на заработной плате водителей.
- более экономичное потребление топлива и использование дорог за счёт централизованного управления транспортным потоком.
- экономия времени, ныне затрачиваемого на управление ТС, позволяет заняться более важными делами или отдохнуть.
- у людей с ослабленным зрением появляется возможность самостоятельно перемещаться на автомобиле.
- минимизация ДТП, человеческих жертв.
- повышение пропускной способности дорог за счёт сужения ширины дорожных полос.

Недостатки:

- Ответственность за нанесение ущерба.
- Утрата возможности самостоятельного вождения автомобиля.
- Надёжность программного обеспечения.
- Отсутствие опыта вождения у водителей в критической ситуации.
- Потеря рабочих мест людьми, чья работа связана с вождением транспортных средств.
- Потеря приватности.
- Минирование беспилотных автомобилей.
- Этический вопрос о наиболее приемлемом числе жертв, аналогичный проблеме вагонетки, стоящий перед компьютером автомобиля при неизбежном столкновении.

Некоторые системы полагаются на инфраструктурные системы (например, встроенные в дорогу или около неё), но более продвинутые технологии позволяют симулировать присутствие человека на уровне принятия решений о рулении и скорости, благодаря набору камер, сенсоров, радаров и систем спутниковой навигации.

V. ВЫВОДЫ

Создание полноценного беспилотного автомобиля — один из самых захватывающих вызовов для технологической мысли начала XXI века для компаний по всему миру. Сегодня беспилотные модели разрабатывают практически все лидеры автоиндустрии — Daimler, Toyota, Volvo, Porsche и даже КамАЗ. Эксперты из немецкой Ассоциации автопрома заявляют, что уже через 15 лет автопилот будет привычным явлением.

VI. СПИСОК ЛИТЕРАТУРЫ

1. O'Toole, Randal. Gridlock: why we're stuck in traffic and what to do about it — Cato Institute, 2009
2. Афанасьев Л. Л. Автомобильные перевозки. — М: Транспорт, 1965.

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ВНЕДРЕНИЯ АСУ ТП НА ХИМИЧЕСКИХ ПРЕДПРИЯТИЯХ В РЕСПУБЛИКЕ БЕЛАРУСЬ

Дорошко А. С.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: doroshko.by@gmail.com

В данной работе проводится анализ результатов внедрения современных автоматизированных систем управления технологическими процессами на химическом производстве, рассматриваются основные проблемы применения современных систем в условиях производства в Беларуси, даются оценки эффективности внедрения таких систем.

ВВЕДЕНИЕ

На сегодняшний день в Республике Беларусь на крупных промышленных предприятиях возникает тенденция к модернизации и замене уже существующих систем управления технологическими процессами. Зачастую, до сих пор работают системы, построенные на базе примитивных схем релейной логики. Такие решения уже давно устарели, так как не отвечают современным требованиям касательно надежности, безопасности, экономической эффективности. Но возникает вопрос об эффективности внедрения современных систем управления.

I. Принципы построения АСУ ТП

На рисунке 1 приведена типовая схема современной АСУ ТП. В большей или меньшей степени все подобные системы на сегодняшний день делают именно таким образом. В первую очередь имеется объект управления. Это может аппарат, совокупность аппаратов, реакторов, трубопроводов, колонн и т.д. В нашем случае под объектом управления будем понимать конкретный технологический процесс. В качестве средств получения информации на химическом производстве используется различного рода измерительная техника. Это различные преобразователи давления, температуры, расхода, плотности и т.д. Они стоят непосредственно на технологических позициях и измеряют необходимые параметры ТП. Так же непосредственно на объекте управления стоят различные исполнительные механизмы, оказывающие непосредственное влияние на протекание процесса. Это могут быть пневмопозиционеры, клапаны, задвижки, насосы, вентиляторы и т.д. В совокупности датчики и исполнительные механизмы являются нижним (первым уровнем) АСУ ТП. В более ранних системах данные с датчиков выводились на центральный пульт управления, где персонал визуально отслеживал показания и принимал исходя из них решение. Так же такие сигналы (как правило дискретные) могли идти на системы примитивной логики, где и принималось решение об

изменении воздействий на объект. Далее сигналы с датчиков поступают на систему локального управления и систему логического управления. Это уже третий уровень АСУ ТП, он же контроллерный. На данном уровне стоят программируемые логические контроллеры. Как видно из рисунка 1, в современных системах оператор так же может оперативно вмешиваться в техпроцесс, но роль человека в последние годы значительно снижается[1].

II. Оценка эффективности внедрения

В целом трудно говорить об общей оценке эффективности внедрения современной АСУ ТП на производстве. В каждом конкретном случае показатели значительно разнятся. Но, как правило, внедрение современных систем позволяет экономить некоторое количество ресурсов, будь то энергоносители или вещества, участвующие в технологических процессах[2]. Экономия может составлять как десятые доли, так и десятки процента. В нашем случае куда более интересна оценка других параметров. В первую очередь это касается безопасности. На данный момент АСУ ТП проектируются исходя из современных требований безопасности, прописанных в стандартах ИЕС 61508. Таким образом значительно повышается промышленная безопасность. Так же внедрение современных систем уменьшает влияние человека на работу установки, что практически полностью исключает вероятность аварии по причине человеческого фактора.

III. Проблемы АСУ ТП в РБ

В Республике Беларусь, как и в любом другом государстве, есть своя специфика внедрения современных систем управления на химическом производстве. Во-первых, немаловажную роль здесь играют особенности белорусского законодательства. Например, ввиду антикоррупционного законодательства на данный момент на всех государственных предприятиях действует тендерная система закупок. Из-за этого предприятие не способно закупить однотипное оборудова-

ние у одного поставщика, что вызывает ряд проблем с унификацией и совместимостью оборудования[3]. Так же резко увеличивается стоимость владения такой системой и возрастают сроки ремонтов, так как становится невозможным держать на месте запасные части и модули, а так же иметь достаточное количество резерва. В результате возникает "зоопарк" приборов, который становится крайне трудно обслуживать. Второй немаловажной проблемой является некомпетентность проектных организаций, разрабатывающих АСУ ТП. К сожалению, часто приходится сталкиваться с некомпетентными проектировщиками, которые не способны самостоятельно разработать рабочую систему управления. Третья проблема – отсутствие в достаточной мере хороших специалистов, одновременно хорошо разбирающихся в технологии производства и проектировании АСУ ТП.

IV. ЗАКЛЮЧЕНИЕ

Таким образом очевидно, что все же внедрение современных АСУ ТП на химических предприятиях в нашей стране дает некоторые результаты, но еще предстоит решать множество проблем для того, чтобы повысить эффективность работы данных систем и улучшить экономические показатели.

V. СПИСОК ЛИТЕРАТУРЫ

1. Хазаров, В. Г. Интегрированные системы управления технологическими процессами / В. Г. Хазаров // СПб.: Профессия, 2009. – 592 с.
2. Олсон, Г. Цифровые системы автоматизации и управления / Г. Олсон // СПб.: Невский Диалект, 2001. – 557 с.
3. Национальный Интернет-портал Республики Беларусь [Электронный ресурс] / Нац. центр правовой информ. Респ. Беларусь. – Минск, 2018. – Режим доступа: <http://www.pravo.by>. – Дата доступа: 15.09.2018.

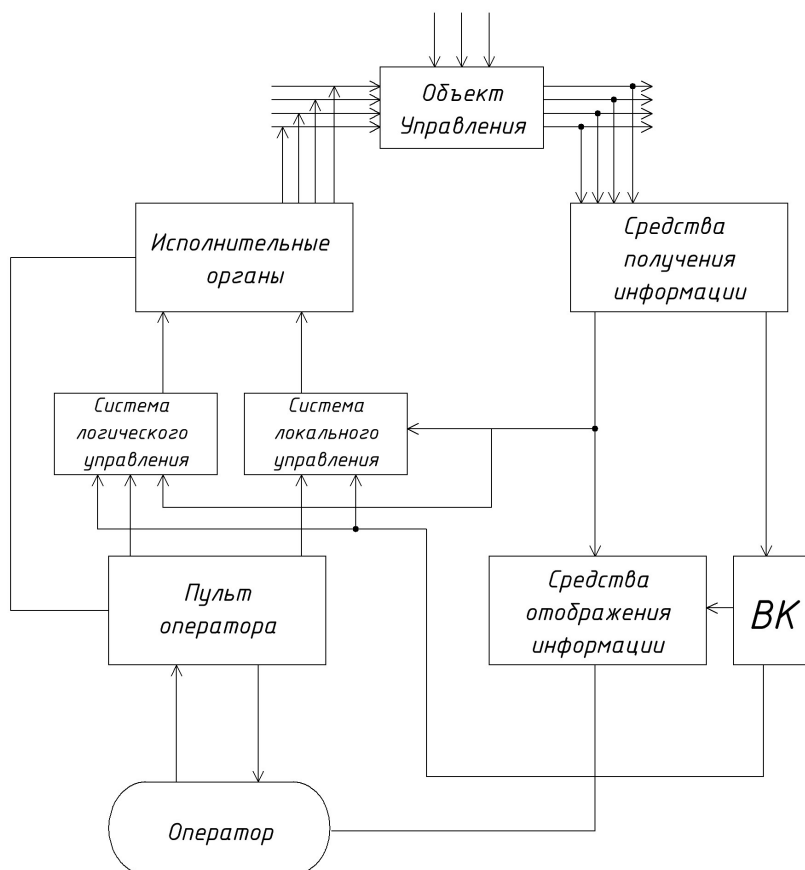


Рис. 1 – Структурная схема АСУ ТП

IMPROVING THE P&O AND INCREMENTAL CONDUCTANCE ALGORITHMS OPERATIONS IN PHOTOVOLTAIC STATION THROUGH A MODEL PREDICTIVE CONTROL

Elzein I. , Petrenko Y. N.

Belarusian National Technical University

Minsk, Belarus

E-mail: Imad.zein@liu.edu.lb, ypetrenko@bntu.by

INTRODUCTION

With availability of fast computers and microprocessors, MPC is increasingly finding application in many other engineering domains such as robotics, automobiles, nuclear and aerospace industries. Major strengths of MPC are abilities to handle multivariable interactions and operating constraints in systematic manner. MPC is formulated as a constrained optimization problem, which is solved on-line repeatedly by carrying out model based forecasting over a moving window of time. More importantly, MPC facilitates optimal control of non-square systems, i.e. systems with unequal number of manipulated inputs and measured outputs. Dynamic models are at the heart of MPC formulations. While initial formulations were based on step response models, a very wide variety of linear / nonlinear black box / mechanistic models are now employed in MPC formulations [1]. This scheme is meant to introduce linear MPC formulation, i.e. one based on linear perturbation models, to a student of advanced control. The linear prediction model for a system can be developed either through linearization of a nonlinear mechanistic model for the system or through identification of a black box linear time series model using input-output data generated by perturbing the system. We assume that the model development exercise is already over and a linear perturbation model for the system under consideration is available with us for controller synthesis.

I. MODEL PREDICTIVE CONTROLLER SCHEME

The ultimate achievement over the next discussions is to improve the operation of the P&O and INC algorithms and this is can be performed through the usage of predicting the future behavior of the desired control variables until a predefined horizon in time. As we utilize the predicted variables, the switching state will be obtained through minimization of a cost function. Predicted behavior of control variables at the next sampling time $k+1$ can be described by a discrete-time set of equations in the actions of a switch is "ON" and when a switch is "OFF".

$$Ipv(k+1) = \frac{T_s}{LV} Vpv(k) + Ipv(k) \quad (1)$$

$$Vc(k+1) = (1 - \frac{T_s}{RC})Vc(k) \quad (2)$$

$$Ipv(k+1) = Ipv(k) - \frac{T_s}{Ln} Vc(k) \quad (3)$$

$$Vc(k+1) = \frac{T_s}{nC} Ipv(k) + (1 - \frac{T_s}{RC})Vc(k) \quad (4)$$

Now after determination of the reference current using the procedure shown in Fig. 1, the cost function can be obtained as following;

$$g_{s=0.1} = |Ipv_{s=0.1}(k+1) - Iref| \quad (5)$$

II. MINIMIZATION OF COST FUNCTION

The objective is to minimize the cost function g . The final switching state for MPPT can be determined using procedure illustrated in Fig. 1. Model Predictive Control (MPC) approach is used for controlling the input current of the DC-DC converter [2]. The designed controller should provide the capability of tracking the reference current generated by the suggested MPPT method with satisfactory dynamic and steady-state performances. Based on the MPC concept, the future behavior of the input current should be predicted separately for each of the two different switching states of the converter using appropriate equations. As discussed, the converter has two different switching states: When S is ON; the voltage of inductor L_1 is:

$$V_{L_1} = V_{pv} = L_1 \frac{di_{L_1}}{dt} \quad (6)$$

In order to make a discrete formation of the continuous differential equations, the below approximation is applied:

$$\frac{dx}{dt} \approx \frac{x[k+1] - x[k]}{T} \quad (7)$$

Then the discrete formation from (6) and (7) is:

$$\frac{i_{L_1}(k+1) - i_{L_1}(k)}{T_s} = \frac{Vpv(k)}{L_1} \quad (8)$$

In which T_s is the controller sampling time. By rewriting (8), the following equation is obtained

for predicting the future behavior of the converter input current in this switching state [3,4]:

$$i_{L_1}(k+1) = \frac{V_{pv}(k)}{L_1} T_s + i_{L_1}(k) \quad (9)$$

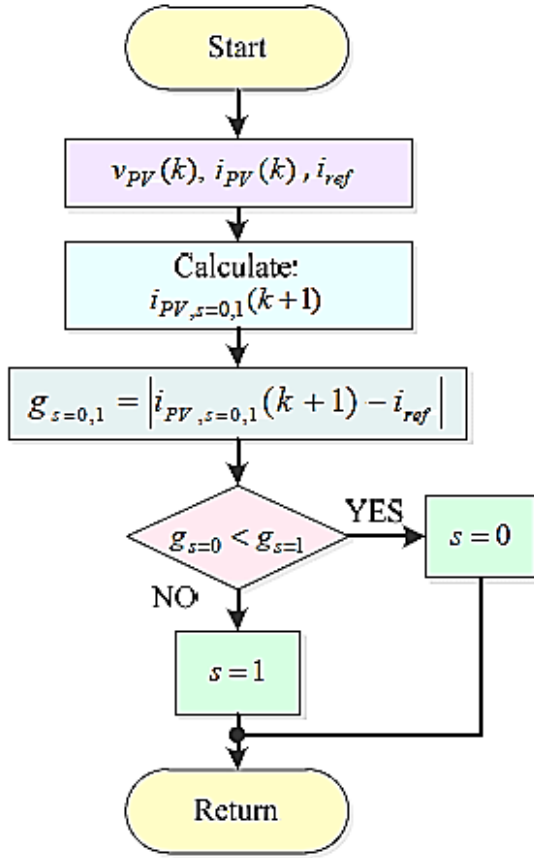


Рис. 1 – Model predictive control process for MPPT

When S is off then for this switching state; the voltage across inductor L_1 is:

$$V_{L_1} = V_{pv} - V_{c1} = L_1 \left(\frac{di_{L_1}}{dt} \right) \quad (10)$$

Thus the discrete form in (10) is:

$$\frac{i_{L_1}(k+1) - i_{L_1}(k)}{T_s} = \frac{V_{pv}(k) - V_{c1}(k)}{L_1} \quad (11)$$

By rearranging formula (11);

$$i_{L_1}(k+1) = \frac{(V_{pv}(k) - V_{c1}(k))T_s}{L_1} + i_{L_1}(k) \quad (12)$$

The next step of designing the controller is to select a suitable cost function based on the control objective :

$$g = |i_{Lref}(k+1) - i_{L_1}(k+1)| \quad (13)$$

The future value of the converter input current is predicted using (9) and (12) in each sampling interval. Then, the switching state that results in minimizing the cost function of (13) is selected to be applied to the converter at next sampling interval. The overall block diagram of the controller is depicted in Fig.2.

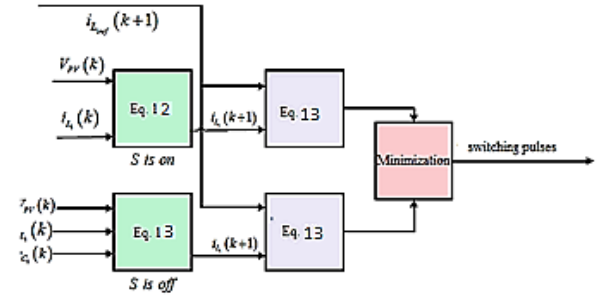


Рис. 2 – Structure model of the model predictive Control

REFERENCES

1. Elzein, I. Fuzzy Logic Controller Design for Photovoltaic System. Information Technologies and Systems 2014 (ITS 2014), Proceedings of the Int. Conf.(BSUIR, Minsk, Belarus, 29 th October, 2014), pp.36-37.
2. Elzein, I. A. A study of maximum power point tracking algorithm for photovoltaic system using a fuzzy logic controller. WIT Transaction on Engineering Sciences. - 2014. - Vol. 96, WIT Press. P.409-419. ISSN: 1743-3533.
3. Imad, Elzein, Yury N. Petrenko. Model predictive control for photovoltaic station maximum power point tracking system. Journal of System Analysis and Applied Information science. Vol. 10, ?4, pp 17 - 25, 2015.
4. Imad Elzein, Yury Petrenko. Adapting Model Predictive Control for a PV Station and Evaluating two different MPPT Algorithms P&O and FLC. Advances in Science, Technology and Engineering Systems Journal (ASTESJ). Volume 2, Issue 3, Special issue on Recent Advances in Engineering Systems. Pp. 741-748, 2017. Walnut, California, USA. ISSN: 2415-6698. DOI: 10.25046/aj020394. pp. 275-280, 2016.
5. I. A. Elzein, Y. N. Petrenko, "An Integration of A Predictive Control Model and MPPT for PV Station". International Conference on Smart Systems and Technologies, Osijek, Croatia. Region 8, IEEE Trans. 2016.

ЭФФЕКТИВНОСТЬ ЛИНЕЙНОГО И НЕЛИНЕЙНОГО 2DOF РЕГУЛЯТОРОВ

Гринюк Д. А., Сухорукова И. Г., Олиферович Н. М., Оробей И. О.

Кафедра автоматизации производственных процессов и электротехники, Белорусский государственный технологический университет
Минск, Республика Беларусь
E-mail: hryniukda@gmail.com

Исследованы различные варианты конструкции регуляторов с двумя степенями свободы (of the two-degree-of-freedom structure, 2DoF). Рассмотрены линейные и нелинейные варианты. В качестве основы настройки и последующего сравнения линейных и нелинейных вариантов 2DoF - регуляторов использованы два варианта интегральных критериев. Показано, что два варианта структур с двумя степенями свободы улучшают работу ПИД-регулятора, но не являются эквивалентными и не обеспечивают одновременной минимизации интегральных критериев возмущения и заданного значения. Среди использованных для комплексной настройки интегральных критериев модифицированный интегральный критерий приводит к более приемлемому для технологических объектов переходному процессу.

ВВЕДЕНИЕ

Качественная настройка промышленных регуляторов способствует повышению стабильности технологических параметров и уменьшает энергозатраты на единицу продукции. Это обусловлено уменьшением переходных процессов и критических отклонений. Практика обучения специалистов теории управления, а также большинство учебной литературы используют наблюдение за переходным процессом вследствие изменения сигнала задания для проверки качества регулирования. В то же время большинство систем стабилизации технологических параметров предназначены для подавления возмущений в процессе функционирования. Иногда возникают задачи, при которых система подвергается частым возмущениям как со стороны канала возмущения, так и ввиду частого изменения сигнала задания.

1. СТРУКТУРЫ 2DOF-РЕГУЛЯТОРОВ

Одним из вариантов решения проблемы является использование two-degree-of-freedom structure (2DoF) [1 – 3]. Существует классический вариант реализации структуры (рис. 1 [2], где SP – сигнал задания; Y_{OUT} – выход; K_P – коэффициент усиления; T_I – время интегрирования; T_D – время дифференцирования; W_A – передаточная функция исполнительного механизма; W_S – передаточная функция измерительного преобразователя; W_O – передаточная функция объекта; F – сигнал возмущения; s – оператор передаточной функции), но можно предложить и альтернативный вариант (рис. 2). Несмотря на введение дополнительных коэффициентов a и b , задачу поиска настроек регулятора можно легко разделить на два этапа: настройка основных коэффициентов K_P , T_I и T_D , а затем поиск дополнительных коэффициентов a и b .

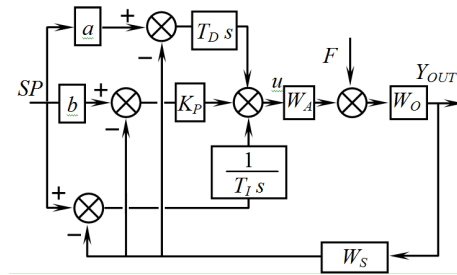


Рис. 1 – Классический вариант регулятора с двумя степенями свободы (КРДСС)

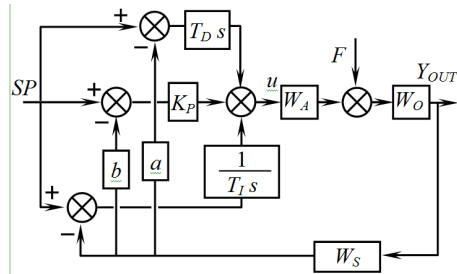


Рис. 2 – Альтернативный вариант регулятора с двумя степенями свободы (АРДСС)

В качестве критерия настройки, как отражено в работе [4], легко можно использовать интегральные критерии [5]. В работе были использованы два интегральных критерия настройки. Первый критерий является классическим, в то же время второй обеспечивает компромиссный вариант между быстродействием, перерегулированием и запасом по устойчивости [5].

$$\int_0^{tf} e(t)^2 dt \rightarrow \min; \quad (1)$$

$$\int_0^{tf} t^2 |e(t)| dt \rightarrow \min. \quad (2)$$

Для исследования была выбрана система со следующими передаточными функциями:

$$W_A(s) = \frac{1}{5s + 1}; W_S(s) = \frac{1}{10s + 1}$$

$$W_A(s) = \frac{1}{(25s + 1)(15s + 1)} \exp(-\tau s)$$

Поскольку транспортное запаздывание времени τ существенно влияет на результат применения интегральных критериев [5], методика настройки была проверена для нескольких значений $\tau = [0; 3; 7,5; 15]$. Минимизация критериев производилась в Matlab. Кроме линейных структур использовался и ПИД-регулятор с нелинейной характеристикой [4,6] При этом нелинейное преобразование использовалось только для интегральной и пропорциональной составляющих регулятора

$$\mu(e) = \text{sign}(e) \ln(1 + |e|) - 1$$

II. РЕЗУЛЬТАТЫ

Анализ результатов показывает, что отличия между оптимальными настройками регуляторов по каналу задания и возмущения могут отличаться в несколько раз. Наибольшее отличие наблюдается во времени интегрирования. Следует отметить, что использование нелинейного преобразования ошибки придает ПИД-регулятору большую универсальность при стабилизации технологических параметров (рис. 3–9, где KID_{SP} – относительное значение интегрального критерия для настроек классического ПИД-регулятора, найденных при настройках по каналу задания; KID_D – относительное значение интегрального критерия для настроек классического ПИД-регулятора, найденных при настройках по каналу возмущения; LR – линейный вариант структуры; NR – нелинейный вариант структуры).

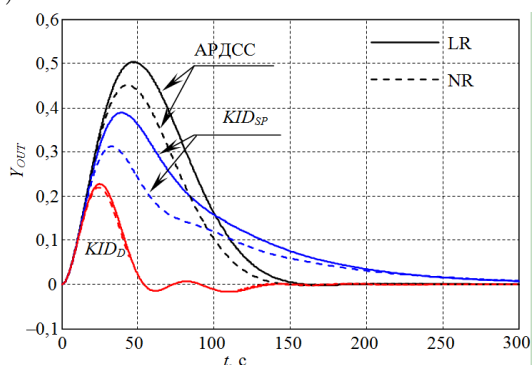


Рис. 3 – Переходной процесс по возмущению для запаздывания $\tau = 0$. Настройка по критерию (2).

Данные показывают, что дополнительные коэффициенты настроек позволяют уменьшить значение интегральных критериев, однако не всегда существенно. Наибольший эффект достигается для объектов без запаздывания или при небольшом его значении. Улучшения наблюдаются также, и по перерегулированию, и по времени переходного процесса. Использование кри-

терия (1), как обычно [5], приводит к высокой колебательности и длительности затухания.

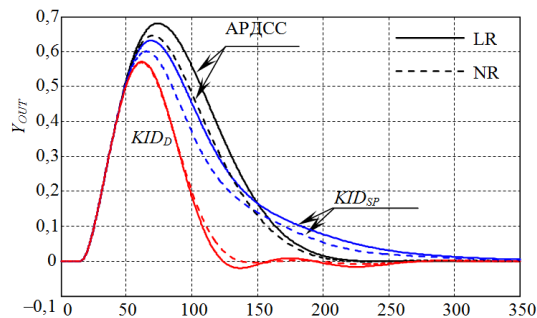


Рис. 4 – Переходной процесс по возмущению для запаздывания $\tau = 0$. Настройка по критерию (2).

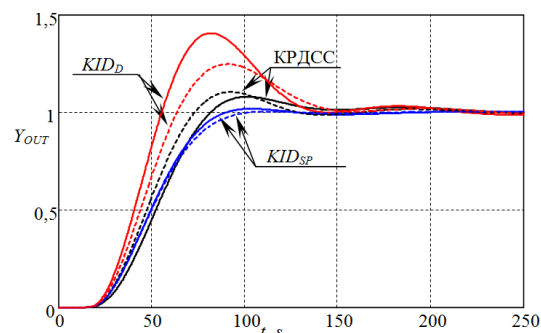


Рис. 5 – Переходной процесс по возмущению для запаздывания $\tau = 0$. Настройка по критерию (2).

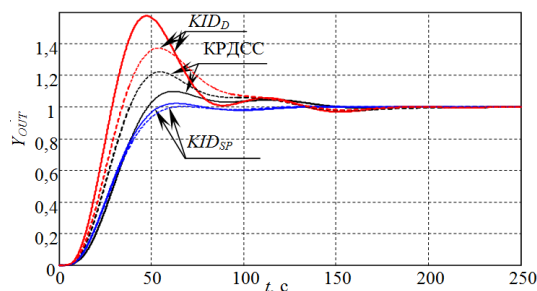


Рис. 6 – Переходной процесс по возмущению для запаздывания $\tau = 0$. Настройка по критерию (2).

1. Åström, K. J., Hägglund, T. Advanced PID control. ISA-The Instrumentation, Systems, and Automation Society; Research Triangle Park, Durham, – 2006. – 461 p.
2. Денисенко, В. В. Компьютерное управление технологическим процессом, экспериментом, оборудованием. М.: Горячая линия – Телеком, 2009. – 608 с.
3. Multi-objective optimization based tuning tool for industrial 2DoF PID controllers [C. Gamboa et al.] // IFAC-PapersOnLine, July 2017. Vol. 50, Issue 1, P.7511–7516.
4. (8) Non-linear PID controller and methods of its setting. [Hryniuk D. et al.] // 2017 Open Conference of Electrical, Electronic and Information Sciences (eStream), 2017, pp.1–4. DOI: 10.1109/eStream.2017.7950327.
5. Модификация интегральных критериев для повышения запаса по устойчивости [Гринюк Д.А. и др.] // Труды БГТУ. 2012. № 6: Физ.-мат. науки и информатика. С. 118–121.
6. Численное исследование алгоритмов уменьшения интегрального насыщения [Гринюк Д.А. и др.] // Труды БГТУ. Сер. VI, Физ.-мат. науки и информатика. 2005. Вып. XIII. С. 140–143.

АВТОМАТИЧЕСКОЕ УПРАВЛЕНИЕ РОБОТАМИ-МАНИПУЛЯТОРАМИ ДЛЯ ДУГОВОЙ СВАРКИ С УЧЕТОМ ТЕХНОЛОГИЧЕСКИХ ОГРАНИЧЕНИЙ

Илюшин И. Э., Кожевников М. М.

Кафедра автоматизации технологических процессов и производств, Могилевский государственный университет продовольствия
Могилев, Республика Беларусь
E-mail: iliue@yandex.by, kmmk@mail.ru

Предложены новые эффективные метод и алгоритм автоматического управления роботами-манипуляторами в процессе дуговой сварки при наличии технологических ограничений. Предложенный подход, в отличие от известных, позволяет эффективно учесть форму сварного шва, а также ограничения, накладываемые на его ориентацию. Эффективность разработанного метода и алгоритма подтверждена результатами экспериментов с моделями промышленных роботов-манипуляторов.

ВВЕДЕНИЕ

В предлагаемой работе рассматриваются задачи управления роботами-манипуляторами (РМ) и роботизированными технологическими комплексами (РТК) в технологическом процессе дуговой сварки (ДС). Стандартный комплекс для ДС (рис.1.) включает манипулятор 1 с шестью степенями свободы (q_1 - q_6) и сварочную горелку 2. Сварная конструкция 3 смонтирована на позиционере 4 с тремя степенями свободы (g_1 - g_3), причем первое от базы сочленение позиционера (g_1) используется для перемещения сварной конструкции в зону выполнения сварочных операций, а оставшиеся два сочленения (g_2 , g_3) – для ориентирования сварных швов [1-3]. Тогда общее число степеней свободы роботизированной сварочной системы – девять.

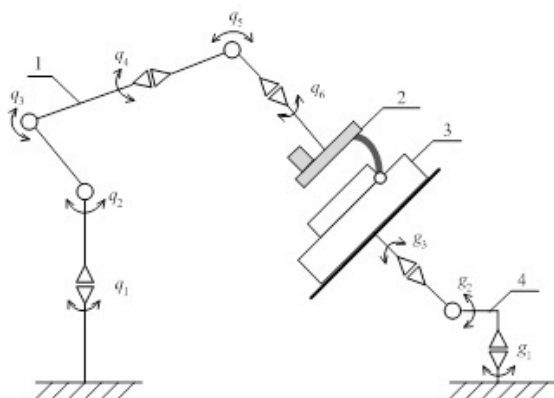


Рис. 1 – Схема РТК для ДС

I. ПОСТАНОВКА ЗАДАЧИ

Ввиду конструктивных особенностей РМ зададим следующие ограничения:

$$\mathbf{q}_{\min} \leq \mathbf{q} \leq \mathbf{q}_{\max}, \quad (1)$$

$$\mathbf{g}_{\min} \leq \mathbf{g} \leq \mathbf{g}_{\max}, \quad (2)$$

где $\mathbf{q} = [q_i]^T$ – конфигурация робота, $\mathbf{g} = [g_i]^T$ – конфигурация позиционера, \mathbf{q}_{\max} , \mathbf{q}_{\min} – максимальное и минимальное ограничения на диапазоны углов в сочленениях РМ, \mathbf{g}_{\max} , \mathbf{g}_{\min} – максимальное и минимальное ограничения на диапазоны углов в сочленениях позиционера.

Ориентация сварного шва определяется углами θ и ε (рис. 2).

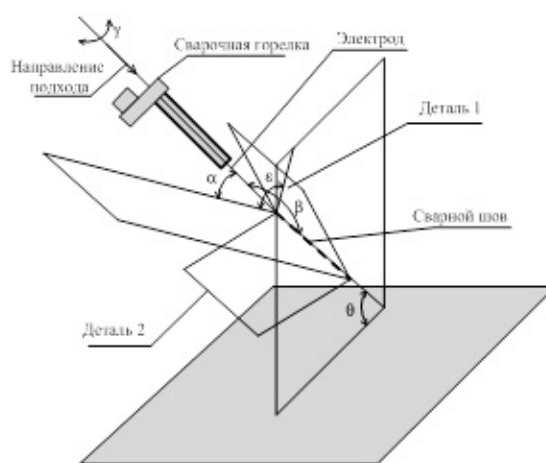


Рис. 2 – Ориентация горелки дуговой сварки и сварного шва

Угол θ – угол между линией сварного шва и горизонтальной плоскостью. Угол ε определяет отклонение биссектрисы угла между свариваемыми деталями от вертикальной плоскости и для значений этих углов определяются допустимые диапазоны изменения. Ориентация горелки дуговой сварки относительно сварного шва задается двумя углами α и β . Угол α характеризует наклон плоскости перемещения сварочной горелки относительно биссектрисы угла образованного свариваемыми деталями, β – угол наклона сварочной горелки к линии сварного шва. Угол γ определяет совместно с α и β ориентацию сварочной горелки.

С учетом технологических ограничений дуговой сварки зададим диапазоны изменения уг-

лов ориентации сварочной горелки и сварного шва:

$$\begin{aligned} \alpha_{min} &\leq \alpha \leq \alpha_{max}, \\ \beta_{min} &\leq \beta \leq \beta_{max}, \\ \gamma_{min} &\leq \gamma \leq \gamma_{max}, \\ \theta_{min} &\leq \theta \leq \theta_{max}, \\ \varepsilon_{min} &\leq \varepsilon \leq \varepsilon_{max}. \end{aligned} \quad (3)$$

Разобьём указанные допустимые диапазоны изменения углов ориентации с параметром дискретизации n . В случае сварки швов сложной формы необходимо изменять ориентацию сварного шва по мере перемещения вдоль него сварочной горелки. Следовательно шов сложной формы разделяется на ряд «простых» сегментов (ограничены точками p_j), ориентирование которых осуществляется индивидуально за счет одновременного перемещения позиционера и робота.

II. АЛГОРИТМ УПРАВЛЕНИЯ

Представим модель конфигурационного пространства манипулятора применив неориентированный граф $DCf=(V,E)$. Вершины графа V описывают конфигурации манипулятора, в которых нет столкновений и которым соответствуют определенные конфигурации позиционера. Ребра графа E определяются как фрагменты траектории горелки между конфигурациями манипулятора из V . Используя данный подход, а также с учетом рассмотренных механических и технологических ограничений разработан алгоритм управления РМ для ДС. Он включает следующие этапы.

1. Используя обратное кинематическое преобразование позиционера, вычисляются углы в его сочленениях, соответствующие дискретным значениям допустимых углов ориентации θ и ε .

2. Используя прямое кинематическое преобразование позиционера, вычисляются множества, определяющие координаты точек сварного шва p_j при конфигурировании позиционера.

3. Используя обратное кинематическое преобразование манипулятора, вычисляются углы в его сочленениях, соответствующие дискретным значениям допустимых углов ориентации сварочной горелки α , β и γ .

4. Каждый элемент множеств q проверяется на принадлежность свободному от столкновений пространству при конфигурировании модели соответствующими углами из множества g . Если элемент принадлежит свободному пространству, то он и соответствующий ему элемент, описывающий углы в сочленениях позиционера, добавляются в множество вершин V , в противном случае элемент отбрасывается.

5. Производится поиск траектории манипулятора между элементами множеств q , включенных в множество V и соответствующим соседним точкам p_j . При существовании таких фрагментов траектории ребро включается в E .

6. Пункты 1-5 выполняются циклически для каждой из точек шва. В результате этого формируется граф, включающий конфигурации манипулятора и позиционера, которые обеспечивают сварку точек шва, и «простые» участки траекторий между этими конфигурациями, движение сварочной горелки вдоль которых не приводит к столкновениям и выходу за технологические и конструктивные ограничения.

7. Ищутся множества возможных маршрутов на графе DCf , образующих траектории манипулятора, согласованные с ориентационными движениями позиционера.

ЗАКЛЮЧЕНИЕ

Эффективность предлагаемого подхода исследована на примере сварочного РТК (рис.3). Роботизированный комплекс обеспечивает сварку металлических пластин и реализован на базе манипулятора IR761, пятиосного позиционера и контроллера RCM3. Результаты тестирования разработанных метода и алгоритма управления показывают, что они позволяют найти траектории движения РМ вдоль сварных швов с эффективным обходом технологических ограничений.

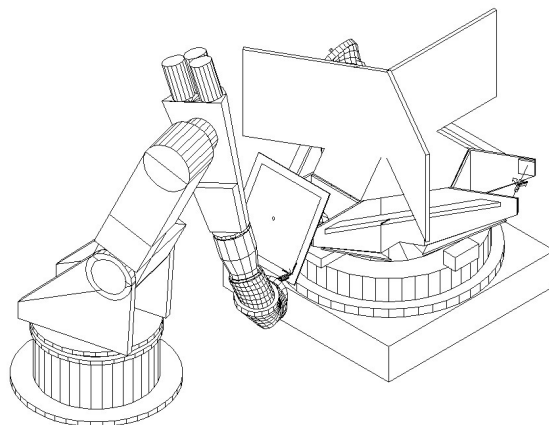


Рис. 3 – Тестовая роботизированная ячейка

1. Pires, J. N. Welding robots: technology, systems issues and applications / J. N. Pires, A. Loureiro, G. Bolmisjo. – Springer-Verlag London Limited, 2006. – 192 p.
2. Климов, А. С. Роботизированные технологические комплексы и технологические линии в сварке: учебное пособие. 2-е изд., испр. и доп. / А. С. Климов, Н. Е. Машнин. – СПб.: «Лань», 2011. – 240 с.
3. Springer Handbook of Robotics / ed. B. Siciliano and O. Khatib. Springer-Verlag, 2016. – 2227 p.

СИСТЕМЫ ПЕРЕМЕЩЕНИЙ В ПРЕЦИЗИОННОМ ОБОРУДОВАНИИ МИКРОЭЛЕКТРОНИКИ

Карпович С. Е., Zubov Г. А., Войтов А. Ю.

Кафедра высшей математики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: mmts@bsuir.by

В работе проанализированы новые тенденции построения современных систем перемещений для прецизионного оборудования нано- и микроэлектроники, предназначенных для создания систем позиционирования, выполнения технологических и измерительных операций, а также представленные в работе подходы могут быть использованы при создании современных обрабатывающих центров и автоматических линий машиностроительного профиля. Они позволяют, в настоящее время, разрешить большинство из проблем, присущих традиционной и широко используемой компоновке на сложившейся элементной базе.

Анализ технологического оборудования производства изделий электронной техники: оптико-механического для генерирования изображений, проекционного совмещения и экспонирования полупроводниковых пластин, контроля изображений; сборочного, включая установки зондового контроля пластин, разделения пластин на кристаллы, посадки кристаллов, монтажа проволочных выводов и других, показывает, что основными узлами, в значительной степени определяющими его производительность, точность, динамику являются устройства координатных перемещений, которые представляют собой автономные механо-аппаратно-программные мехатронные комплексы, предназначенные для осуществления рабочих, установочных и юстировочных перемещений кристаллов, полупроводниковых пластин, инструментов и элементов проекционной оптической системы.

Переход на меньшую топологическую норму, увеличение диаметра кремниевых пластин требует серьезной технологической подготовки производства не только у основных производителей изделий микроэлектроники на финишной стадии, но и у поставщиков материалов и компонентов. Изготовление кремниевых пластин с требуемой плоскостностью при увеличенных габаритных размерах представляет серьезную проблему. Нужно более точно выдерживать толщину пленки фоторезиста при покрытии кремниевой пластины, более точно дозировать и имплантировать легирующие примеси, более точно выполнять совмещение изображений элементов топологии фотошаблонов различных слоев при мультиплицировании топологии кристалла СБИС по кремниевой пластине, более точно позиционировать зонды на уменьшенные контактные площадки интегральной схемы при межоперационном зондовом контроле, более точно позиционировать режущий инструмент на большей протяженности трассы реза при сокращении расстояния между трассами и, наконец, все уз-

лы оборудования должны поддерживать субмикронное разрешение и повторяемость всех технологических процессов, обеспечивающих приемлемый выход годных кристаллов с одной кремниевой пластины.

Применение механизмов параллельной кинематики в качестве исполнительных устройств систем пространственных перемещений широкого назначения, включая системы координатных совмещений прецизионного оборудования микро- и нанoeлектроники, системы позиционирования, технологических и измерительных операций в трехмерном пространстве современных обрабатывающих центров и автоматических линий машиностроительного профиля, позволяет в настоящее время разрешить большинство из проблем, присущих традиционной и широко используемой компоновке и элементной базе робототехнических комплексов в составе станков и промышленных роботов.

Сравнительный анализ многокоординатных систем перемещений на линейных шаговых двигателях и систем перемещений, конфигурируемых из предложенных механизмов параллельной кинематики, управляемых поворотными, сегментными, линейными и планарными приводами прямого действия, показал, что дальнейшее повышение в 2...3 раза показателей назначения, таких как точность и быстродействие в новом поколении оборудования, возможно только путем внедрения систем перемещений на механизмах параллельной кинематики, в которых прямое электромагнитное преобразование движения реализуется через параллельные механические структуры.

Наибольший эффект в прецизионном оборудовании дают механизмы параллельной кинематики при использовании в качестве привода синхронных двигателей с возбуждением от постоянных магнитов. При этом необходимо отметить, что механизмы параллельной кинематики до шести степеней свободы включительно выгля-

дятся наиболее приемлемыми с точки зрения построения на их основе перспективных координатных систем, способных реализовывать практически весь спектр требуемых движений инструмента и заготовки в трехмерном пространстве. Но, несмотря на все преимущества и свойства таких механизмов, приходится констатировать, что в настоящее время их применение в автоматическом оборудовании и робототехнике пока ограничено из-за сложности управления параллельными механизмами и недостаточной проработанности алгоритмов и методик, которые позволяли бы получать как оптимальные варианты конструкций, так и наиболее подходящие для выбранной конструкции алгоритмы автоматического управления приводами.

Представленные в докладе подходы, математические модели, алгоритмы и программы компьютерного моделирования кинематики и динамики исполнительных механизмов параллельной кинематики с шестью степенями свободы на поворотных, сегментных и планарных шаговых двигателях направлены на оптимизацию структурно-кинематических схем систем перемещений, построенных на этой элементной базе, а также должны обеспечить разработку быстрых алгоритмов регулирования и управления системами приводов или многокоординатными двигателями для реализации требуемых траекторных задач с заданными характеристиками точности и быстродействия. Разработанные алгоритмы и программное обеспечение решения прямой и обратной задач кинематики разработанных базовых механизмов могут быть рекомендованы для компьютерного формирования рабочей области, калибровки возможных траекторий по кинематическим, динамическим и точностным показателям, необходимым для обеспечения условий реализации соответствующих технологических операций в технологическом оборудовании.

В работе предложена имитационная модель динамики в виде модульного описания механической структуры с возможностью автоматического импортирования структурно-параметрического представления механизмов из современных САПР (CAD-систем), таких как Pro/ENGINEER, SolidWorks и др. в среду MATLAB/Simulink, позволяющая автоматизировать решение прямой и обратной задач динамики с интерактивной визуализацией результатов.

Исполнительные механизмы параллельной кинематики на поворотных и планарных шаговых двигателях могут быть рекомендованы для использования в качестве технологических и транспортных систем при разработке и создании автоматических линий по производству прецизионных деталей и изделий в приборостроении, а также в производстве изделий общемашиностроительного назначения. Системы перемещений на кольцевом сегментном многокоординатном приводе рекомендуется использовать в оптико-механическом оборудовании микроэлектроники при создании адаптивных проекционных систем с программной коррекцией положения оптических элементов в процессе сборки и эксплуатации. Это, в первую очередь, оборудование с проекционными системами нанометрового разрешения, в которых важнейшими являются точность позиционирования элементов и стабильность поддержания геометрических параметров в процессе эксплуатации.

1. Карпович, С.Е. Системы многокоординатных перемещений и исполнительные механизмы для прецизионного технологического оборудования : моногр. / С.Е. Карпович [и др.] ; под. ред. д-ра техн. наук, проф. С.Е. Карповича. – Минск : Бестпринт, 2013. – 208с.
2. Моделирование механизмов параллельной кинематики в среде MATLAB/Simulink : моногр. / С.Е. Карпович [и др.]. – Минск : Бестпринт, 2013. – 153 с.
3. Аваков, С.М. Математическая модель исследования колебательного движения каретки координатного стола / С.М. Аваков, В.В. Жарский, С.Е. Карпович // Инженерный вестник. – 2006. – № 1(21)/5. – С. 40–45.
4. Дайняк, И.В. Построение уравнений движения для многокоординатных систем автоматизированного оборудования / И.В. Дайняк // Машиностроение : сб. науч. тр. – Вып. 18 – Минск : Технопринт, 2002. – С. 290–293.
5. Синхронный стержневой мотор: пат. 4760 Респ. Беларусь, С1 / В.В. Жарский, В.И. Лозовский, А.И. Михалев, Н.К. Трусов; заявитель В.В. Жарский. – № а 20000146; заявл. 16.02.2000; опубл. 30.12.2002 // Офиц. бюллетень / Нац. центр. интеллект. собственности. – 2002. – № 4 – С. 174.
6. Управление электромеханическим приводом прямого действия на основе методов голономных автоматических систем / В.В. Жарский [и др.]. // Современные средства связи : материалы 12-й Междунар. науч.-техн. конф., Минск, Респ. Беларусь, 2007 г. – Минск : ВГКС, 2007. – С. 100–101.
7. Wicker, J. Modeling of small fluctuations in a conservative link: Dynamic behaviour of spatial linkages / J. Wicker // Trans. of the ASME, B., 1979. – Vol. 9. – Nr. 1. – P. 266–267.

МОДЕЛЬ ОНТОЛОГИИ ПОНЯТИЙНЫХ КАТЕГОРИЙ ПЛАНИМЕТРИИ

Корчажкина О. М.

Институт кибернетики и образовательной информатики

Федерального исследовательского центра «Информатика и управление» Российской академии наук

Москва, Российская Федерация

E-mail: {olgakomax}@gmail.com

В статье обсуждается модель онтологии, содержащей основные понятия из раздела «Планиметрия» с точки зрения конструктивных функций предметной области «Геометрия». Приводится пример решения конкретной задачи на построение, представляющей понятие «геометрическое место точек», которое может быть включено в базу знаний соответствующей экспертной системы.

ВВЕДЕНИЕ

Несмотря на то, что знания, представленные в школьном курсе геометрии, довольно легко поддаются формализации, проблема создания экспертных систем с последующим использованием средств ИКТ для решения творческих – неформализованных – геометрических задач на построение вызывает определённые трудности. Известно, что одним из необходимых компонентов процедуры разработки любой экспертной системы является этап концептуализации, на котором «проводится содержательный анализ проблемной области, выявляются используемые понятия и их взаимосвязи, определяются методы решения задач» [1, с. 37]. Поэтому онтология понятий из предметной области «Планиметрия», предоставляющая знаниевые инструменты для многочисленных построений, выдвижения гипотез и их верификации, может представлять определённый интерес для разработчиков интеллектуальных систем.

I. АРХИТЕКТУРА ОНТОЛОГИИ РАЗДЕЛА «ПЛАНИМЕТРИЯ»

Обобщённые понятия в курсе планиметрии подразделяются на несколько уровней. Прежде всего, это сущностные категории: абстрактные сущностные категории – теорема, аксиома, формула, вывод, правило, признак, доказательство и деятельностные (функциональные) сущностные категории – равенство, подобие, симметрия, параллельность, система координат, пересечение, область, предел, степень, геометрическое место и пр. На следующем уровне обобщения находятся фундаментальные геометрические объекты: базовые геометрические элементы – точка, прямая, кривая, плоскость и геометрические фигуры – треугольник, окружность, квадрат, трапеция и пр. Затем идёт уровень производных геометрических объектов: производные простые элементы – луч, отрезок, дуга; производные объекты, возникающие как соотношение базовых и простых элементов между собой – параллельные прямые, перпендикуляр, угол, биссектриса; производные объекты, возникающие как соотношение между

простыми элементами и фигурами, как правило, поддающиеся измерению – сторона, внутренний/внешний угол, основание, вершина, высота, медиана, радиус, диаметр, хорда, площадь круга/ромба, длина окружности, периметр, сектор, сегмент, средняя линия и пр.

Существует ещё три группы универсальных понятий: линейные и угловые единицы измерения – сантиметр, метр, градус, радиан, константы – число $\pi = 3,14159$, «золотое сечение» 1.618, константа окружности 6,28308 и конкретные правила (пропорции) и формулы – правило «золотого сечения», формулы для вычисления углов, площадей, длин отрезков, сторон, дуг, окружностей, соотношений между углами в треугольнике и пр. Можно считать, что универсальные понятия в планиметрии находятся над иерархией описанных выше категорий, поскольку они конкретизируют их, то есть «вдыхают жизнь» в неопределённые понятия.

В соответствии с приведённой классификацией можно составить простейшую онтологию понятийных категорий планиметрии с учётом функциональных возможностей конструктивной геометрии, то есть разветвлённую сеть концептов, находящихся во взаимодействии друг с другом на разных уровнях обобщения:

сущностные категории

- абстрактные
- деятельностные (функциональные)

фундаментальные геометрические объекты

- базовые геометрические элементы
- геометрические фигуры

производные геометрические объекты

- производные простые элементы
- производные объекты, возникающие как соотношение базовых и простых элементов между собой
- производные объекты, возникающие как соотношение между простыми элементами и фигурами

универсальные понятия

- линейные и угловые единицы измерения
- константы

- конкретные правила (пропорции) и формулы

II. «ГЕОМЕТРИЧЕСКОЕ МЕСТО ТОЧЕК» КАК СУЩНОСТНО-ДЕЯТЕЛЬНОСТНАЯ КАТЕГОРИЯ ПЛАНИМЕТРИИ

Особое место среди сущностно-деятельных категорий занимают обобщённые понятия, которые, с одной стороны, заключают в себе свойства целого класса однородных или сходных геометрических объектов, а с другой – сами могут являться реальными геометрическими объектами, образованными упорядоченной совокупностью определённых элементов как части объектов данного класса. К подобным категориям относится распространённое понятие геометрическое место точек (ГМТ). Характеризуя заданные свойства совокупности точек, принадлежащих некоей группе геометрических объектов, любое ГМТ – это геометрическая фигура, состоящая из всех точек, удовлетворяющих этому свойству. Если находится ещё какая-либо точка, обладающая этим свойством, то она определённо принадлежит указанному ГМТ. И наоборот: если точка принадлежит указанному ГМТ, то она определённо обладает свойством, которым обладают все остальные точки этого ГМТ.

III. ПРИМЕР КОМПОЗИЦИИ ПРЕОБРАЗОВАНИЙ ДЛЯ ОПРЕДЕЛЕНИЯ ГМТ

В качестве примера рассмотрим построение ГМТ – середин хорд, исходящих из общей точки окружности (рис. 1). Эта задача является промежуточным этапом одного из решений парадокса Бертрана [2, с. 36]. Искомым ГМТ является окружность, внутренняя по отношению к данной окружности, касающаяся её в точке исхода хорд и имеющая радиус, равный половине радиуса данной окружности. Алгоритм решения задачи состоит в следующем:

1. Строится окружность с центром в т.О.
2. На окружности выбирается произвольная точка А, из которой строится веер хорд: АВ₁, АВ₂, АВ₃, АВ₄, ... АВ_n.
3. Находится середина каждой хорды (точки С₁, С₂, С₃, С₄, ... С_n), которые соединяются замкнутой кривой.
4. Доказывается, что полученное ГМТ является окружностью, проходящей через точки А и О и имеющей радиус вдвое меньший, чем исходная окружность [3].

Композиция элементарных преобразований для нахождения середин отрезков, которые в данной задаче являются хордами исходной окружности, описана в [4, с. 133]. Для построения композиции преобразований всей задачи требуется осуществить поворот хорды в пределах развёрнутого угла k с вершиной в точке А, причём длины хорд будут ограничиваться исходной окружностью, изменяясь от нуля до её диаметра АВ₃ при 90 градусам к сторонам развёрнутого угла и да-

лее снова до нуля. То есть второй конец каждой хорды будет являться точкой пересечения исходной окружности и луча, исходящего из точки А. Следовательно, композиция преобразований, изображённая на рис. 4 в [4, с. 133], может быть дополнена решением, выполненным на формальном языке для одной из хорд АВ_n и представленным на рис. 2.

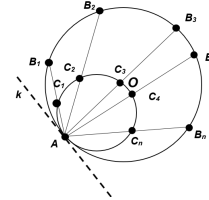


Рис. 1 – Задача на поиск ГМТ середин хорд, исходящих из точки окружности

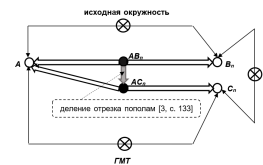


Рис. 2 – Упрощённая композиция преобразований на поиск ГМТ середин хорд, исходящих из точки окружности

IV. ЗАКЛЮЧЕНИЕ

Предложенная модель онтологии понятийных категорий планиметрии позволяет совместить знаниевые и прагматические компоненты предметной области «Геометрия» с целью поиска новых способов решения задач на построение. На примере понятия «геометрическое место точек» было продемонстрировано решение нестандартной задачи из области конструктивной геометрии, требующей многочисленных построений при решении традиционными способами. Использование формализованного языка и функциональных возможностей обобщённых понятий делают возможным создание экспертных систем, с помощью которых упрощается решение творческих задач на построение.

СПИСОК ЛИТЕРАТУРЫ

1. Ясницкий, Л. Н. Интеллектуальные системы : учебник / Л. Н. Ясницкий. – М.: Лаборатория знаний, 2016. – 221 с.
2. Корчажкина, О. М. Метод виртуального статистического эксперимента при решении вероятностных задач // Информатика и образование. – 2018. – № 6. – С. 36–41.
3. Корчажкина, О. М. Решение геометрических задач с применением методики Поля и привлечением ИКТ // Электронное приложение к журналу «Математика в школе». – 2018. – № 1.
4. Межень, А. Л., Пашкевич, Е. С. Сафоненко, К. А. Онтология предметной области геометрических преобразований // Информационные технологии и системы 2017 (ИТС 2017): материалы международной научной конференции, БГУИР, – Минск, Беларусь, 25 октября 2017 г. / редкол. Л.Ю. Шилин [и др.]. – Минск: БГУИР, . –2017. –305. с. . – С. 132-133.

ИНВАРИАНТНАЯ СИСТЕМА АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ УРОВНЕМ ВОДЫ В БАРАБАНЕ ПАРОВОГО КОТЛА БЕЗ ИЗМЕРЕНИЯ ВНЕШНИХ ВОЗМУЩЕНИЙ

Кулаков Г. Т., Кухоренко А. Н.

Белорусский национальный технический университет

Научно-исследовательский институт пожарной безопасности и проблем чрезвычайных ситуаций МЧС

Республики Беларусь

Минск, Республика Беларусь

E-mail: ank-17@yandex.by

Использование передаточной функции оптимального регулятора позволяет произвести настройку системы автоматического управления по требуемым показателям качества. В этом случае в качестве критерия качества используют заданную передаточную функцию системы, структура которой зависит от структуры передаточной функции объекта управления, требуемой скорости изменения регулируемого параметра и максимально допустимой величины регулирующего воздействия при отработке скачка задания. Существенное улучшение качества регулирования достигается использованием комбинированного принципа регулирования по отклонению и возмущению как с измерением наиболее опасного внутреннего возмущения или крайнего внешнего возмущения, так и с косвенным выделением эквивалентного внешнего возмущения.

ВВЕДЕНИЕ

Тепловые электростанции составляют основу энергетики Республики Беларусь. В настоящее время на ТЭЦ Республики Беларусь установлено 234 барабанных котла. Реализация контрактного соглашения между ГУ «Дирекция строительства атомной электростанции» и ЗАО «Атомстрой-экспорт» (Российская Федерация) о строительстве двух энергоблоков атомной электростанции увеличит установленную мощностью энергосистемы республики на 2400 МВт. При этом для каждого энергоблока АЭС проектом установлены четыре барабанных парогенератора, которые будут работать в базовой части графиков электрических нагрузок энергосистемы, а пиковую и полупиковую части нагрузок будут покрывать тепловые электростанции, способные работать в широком диапазоне изменения нагрузок. Наибольшее влияние на надежность и экономичность работы барабанных паровых котлов ТЭС и парогенераторов АЭС оказывают системы автоматического управления (далее – САУ) уровнем воды в барабане, которые работают совместно с технологическими защитами по повышению или упуску воды. Как перепитка, так и упуск воды в барабане приводят к аварийному останову котла. Вместе с тем существующие САУ уровнем воды в барабане парового котла или парогенератора и их модификации, получившие широкое распространение на ТЭС и АЭС, на практике при глубоких изменениях нагрузки, оказались недостаточно эффективными. Это снижет безопасность, надежность, экономичность и долговечность работы оборудования ТЭС и АЭС.

На уровень воды в барабане котла влияет множество возмущающих воздействий. При этом можно выделить основные:

- изменение расхода питательной воды;
- изменение температуры питательной воды;
- изменение расхода перегретого пара;
- изменение давления внутри барабана котла;
- изменение расхода топлива.

Изменение расхода питательной воды взаимосвязано с изменением нагрузки потребителя. При увеличении расхода пара увеличивается расход питательной воды и наоборот. Поэтому наибольшее влияние на отклонение уровня воды в барабане котла оказывают глубокие изменения расхода перегретого пара. Качество переходных процессов изменения уровня воды в барабане котла во многом зависит от выбора структуры системы регулирования, закона регулирования и оптимальных параметров динамической настройки регулятора. Применение метода структурно-параметрической оптимизации и теории инвариантности позволяют за счет изменения структуры системы и оптимизации динамической настройки системы автоматического управления существенно улучшить качество поддержания уровня воды в барабане котла [1]. Преимуществом предложенной системы управления является возможность замены отдельных контуров системы и всей многоконтурной САУ передаточными функциями критериев качества, что позволяет снизить порядок дифференциального уравнения, описывающей динамику многоконтурной САУ, а также один параметр динамической настройки системы $T_{зд}$ в каждом контуре регулирования [2–3] (Рис 1.).

I. СОДЕРЖАНИЕ ДОКУМЕНТА

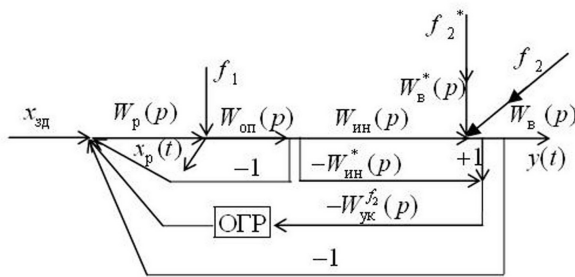


Рис. 1 – Инвариантная САУ уровнем воды в барабане котла

При этом численное значение параметра динамической настройки $T_{зд}$ в передаточной функции критерия качества позволяет экспресс-методом определять прямые показатели качества переходного процесса без построения переходного процесса в САУ при обработке задающего сигнала. Кроме того при обработке задания отсутствует перерегулирование, а уменьшение численного значения $T_{зд}$ позволяет не только повысить быстродействие системы, но и уменьшить максимальную динамическую ошибку регулирования при обработке внутреннего возмущения, а также учесть максимальную допустимую величину регулирующего воздействия при расчете оптимального значения $T_{зд}$. Оптимальная структура САУ формируется таким образом, чтобы каждое воздействие ($x_{зд}$, f_1 и f_2) обрабатывалось по своему каналу. Так, например, для качественной обработки внешнего возмущения без его измерения, параллельно реальному объекту по каналу регулирующего воздействия формируют его модель с адекватной динамикой, а разность между регулируемой величиной и выходом модели подают на вход устройства компенсации эквивалентного внешнего возмущения. Особенностью теплоэнергетических объектов яв-

ляется зависимость их динамических характеристик от уровня нагрузки котла. Поэтому для сохранения высокого качества регулирования в широком диапазоне изменения нагрузок заданную передаточную функцию системы всех регулируемых устройств и внутренних моделей корректируют в зависимости от уровня нагрузки по нелинейным зависимостям. Это позволяет сохранить высокую динамическую точность САУ при основных воздействиях во всем диапазоне изменения нагрузок котла. При этом для участков без запаздывания достигается полная инвариантность при обработке наиболее опасного внутреннего возмущения, а при изменении внешних возмущений – инвариантность с точностью до ε . Это позволяет при модернизации систем управления теплоэнергетическими процессами уменьшить как капитальные, так и эксплуатационные затраты, повысить уровень безопасности, надежности, долговечности, экономичности и экологичности работы ТЭС и АЭС в переменных режимах работы.

II. СПИСОК ЛИТЕРАТУРЫ

1. Кулаков, Г. Т. Оптимизация переходных процессов изменения уровня воды в барабане паровых котлов / Г. Т. Кулаков, А. Н. Кухоренко // Энергетика – Изв. высш. учеб. заведений и энерг. объединений СНГ. – 2014. – № 4. – С. 63–74.
2. Кулаков, Г. Т. Усовершенствование комбинированной САУ уровня воды в барабане парового котла / Г. Т. Кулаков, А. Н. Кухоренко // Материалы 15-й Международной научно-технической конференции «Наука – образованию, производству, экономике»: В 4-х т. Т. 1. – Минск: БНТУ, 2017. – С. 125.
3. Регулятор уровня воды в барабане парогенератора: пат. 027537 Евразийское патентное ведомство, F22D 5/30 (2006 01) / Г. Т. Кулаков, А. Т. Кулаков, А. Н. Кухоренко; заявитель Белорусский национальный технический университет; заявл. 22.04.2015; опубл. 31.08.2017.

МЕТОДИКА ОЦЕНКИ И ПРОГРАММНОЙ КОРРЕКЦИИ НЕРАВНОМЕРНОСТИ АЧХ ВЧ-ТРАКТА ГЕНЕРАТОРА ДЛЯ ШИРОКОПОЛОСНОГО ПЕРЕСТРАИВАЕМОГО АКУСТООПТИЧЕСКОГО ФИЛЬТРА

Давыдов М. В., Ляшук Ю. А., Мальцев О. С., Соколовский Д. В.
Центр 1.6, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: maltsev@bsuir.by

В докладе предложена структурная схема микропроцессорной системы управления параметрами широкополосного перестраиваемого акустооптического фильтра, а также методика программной коррекции неравномерности АЧХ ВЧ-тракта данной системы с целью повышения эффективности управления параметрами фильтров.

ВВЕДЕНИЕ

Одним из перспективных направлений повышения эффективности спектроскопических исследований биообъектов является разработка широкополосных перестраиваемых акустооптических фильтров (ШПАОФ). Оптическая фильтрация широко используется для обнаружения и идентификации микрообъектов и для визуализации их структуры. Фильтр настраивается посредством подачи ВЧ-сигнала заданной мощности и частоты на выделение участка спектра излучения, который определяется характерными свойствами исследуемого объекта или его элементов, что позволяет получать изображения, контрастно отображающие именно эти свойства [1]. Основными характеристиками ШПАОФ являются: рабочий оптический диапазон и связанный с ним диапазон ВЧ-частот, а так же дифракционная эффективность фильтра - процент мощности спектральной составляющей оптического излучения заданной длины волны, который фильтр может дифрагировать из проходящего через него оптического излучения. Дифракционная эффективность зависит от множества факторов, одним из определяющих является величина ВЧ-мощности, передаваемая в фильтр. Таким образом, наличие широкополосного ВЧ-генератора с высокой дискретизацией перестройки частоты и выходной мощности является одним из требований по обеспечению качественного функционирования ШПАОФ.

1. СТРУКТУРА СИСТЕМЫ УПРАВЛЕНИЯ

В общем виде структурная схема (рис. 1) содержит цифровой синтезатор частот (ЦСЧ), аттенюатор (Атт), усилительный каскад (УК), каскад фильтров (КФ), каскад оценки согласования и уровня мощности (КО).

Цифровой синтезатор частот обобщенное название многокомпонентного каскада, основным элементом которого является цифровой синтезатор прямого синтеза частот (ЦСПСЧ), обеспечивающий формирование сигналов различных форм во всей требуемой полосе частот.

Для решения задач управления параметрами ШПАОФ может применяться AD9959 [2]. Также, для обеспечения предварительного усиления и компенсации затуханий в тракте, устранения побочных спектральных компонентов в выходном сигнале, в состав каскада могут быть включены предварительный усилитель и фильтр нижних частот. Для прецизионной установки выходной результирующей мощности на выходе ЦСЧ, могут использоваться различные программно-управляемые аттенюаторы различной битности.



Рис. 1 – Структурная схема микропроцессорной системы управления

ШПАОФ рассматриваемого типа обычно возбуждается в полосе частот порядка 10-200 МГц, при мощности, как правило, не превышающей нескольких ватт [2]. Поэтому, для обеспечения требуемого уровня выходной мощности, используется усилительный каскад, который может состоять из одного или нескольких ВЧ-усилителей данного диапазона.

Для подавления высших гармоник в спектре выходного сигнала, возникающей из-за нелинейности УК, используется каскад полуктавных ФНЧ, которые обеспечивают перекрытие всего рабочего диапазона частот.

Для введения контроля соответствия уровня выходной мощности заданному, разработан КО, который состоит из двунаправленного ответвителя мощности и детектора ВЧ-мощности. Общее управление всеми функциональными узлами осуществляется при помощи ПО микропроцессора.

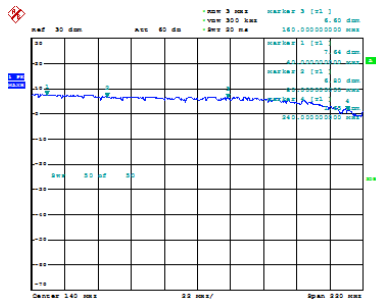


Рис. 2 – АЧХ ВЧ-тракта без коррекции

В результате измерений параметров разработанного по данной схеме устройства было определено, что отклонения АЧХ (рис. 2) могут составлять до 2дБ, что значительно ухудшает точность управления параметрами ШПАОФ, для выравнивания характеристики была разработана методика программной коррекции АЧХ.

II. МЕТОДИКА ОЦЕНКИ И КОРРЕКЦИИ АЧХ ВЧ-ТРАКТА

Неравномерность АЧХ объясняется нелинейностью коэффициента усиления усилительных каскадов, разбросом их характеристик, различием режимов работы, неравномерностью характеристик фильтров и аттенюатора, затуханиями в кабельных сборках.

Методика программной коррекции содержит три этапа. Для реализации первого этапа при помощи программно-аппаратного комплекса, состоящего из ПО управления микропроцессорной системой и анализатора спектра, формируется график величин выходной мощности при изменении как выходной частоты, так и мощности. При этом, в зависимости от сложности системы, методика позволяет оценить влияние как отдельного элемента/каскада — например, аттенюатора, так и совокупности - «выходная амплитуда ЦСПСЧ - аттенюатор». Это позволяет оценить неравномерность АЧХ на всем возможном массиве управляющих команд, определить уровни мощности при которых усилительные каскады переходят в состояние компрессии (рис. 3).

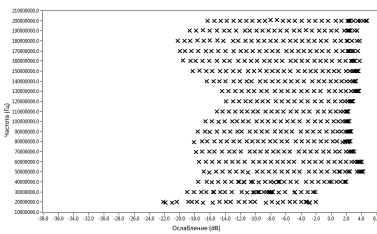


Рис. 3 – Величина выходной мощности в зависимости от частоты при переборе всех возможных значений управляющих команд

На втором этапе полученные данные связывают реальные значения выходной мощности с теми, что были заданы программно. Основываясь на этом, на третьем этапе формируется оптимальный набор состояний всех управляющих команд и выполняется запись значений в блок управления, с учетом которых производится подстройка всех каскадов.

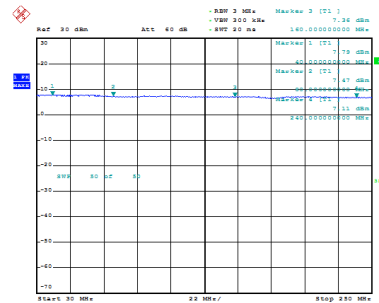


Рис. 4 – АЧХ ВЧ-тракта с коррекцией

Результат работы ПО системы управления представлен на рис. 4 — при выходной мощности, равной 8 дБм (с учетом усредненных потерь в кабеле порядка 0.6 дБ) — неравномерность в полосе рабочих частот 0.5дБ.

III. Выводы

Предложенная структура микропроцессорной системы управления соответствует всем современным тенденциям в области разработки современных ВЧ-устройств. Применение микропроцессора как управляющего элемента позволяет гибко управлять корректировать негативное влияние как отдельных элементов/каскадов, так и минимизировать и учесть взаимовлияние нескольких. Основываясь на этом, представлена методика, позволяющая качественно оценить и скорректировать это влияние.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Перестраиваемый акустооптический фильтр на основе кристаллов парателлурита /А. И. Колесников, И. А. Каплунов, В. Я. Молчанов, С. Е. Ильяшенко, Р. М. Гречишкин // – ИЗВ. ВУЗОВ. ПРИБОРОСТРОЕНИЕ –2012. –т.55, –№10. –С.73-77
2. Кандырин, Н. П.Формирователь многочастотного сигнала на основе прямого цифрового синтезатора для управления акустооптическими фильтрами / Н. П. Кандырин. – Харьков.: Збірник наукових праць Харківського університету Повітряних Сил, 2016. – Вип.№1. – С.70-75.

МЕТОД БИОЭЛЕКТРОГРАФИИ НА СОВРЕМЕННОМ ЭТАПЕ

Новицкий И. О., Шилин Л. Ю.

Кафедра информационных технологий и автоматизированных систем. Белорусский государственный университет информатики и радиоэлектроники

Минск, республика Беларусь

E-mail: Email: mijikai1994@gmail.com

Диагностика физического состояния здоровья пациента на основе обработки и анализа изображений газоразрядного свечения пальцев руки в поле высокой напряженности.

ВВЕДЕНИЕ

Эффект Кирлиана, или эффект газоразрядной визуализации — коронный барьерный разряд в газе. Объект предварительно помещается в переменное электрическое поле высокой частоты, при котором возникает разность потенциалов между электродом и исследуемым объектом от 5 до 30 кВ. Эффект, подобно статическому разряду или молниям, наблюдается на биологических объектах, а также на неорганических образцах разного характера. Исследования, осуществленные С.Д. и В.Х. Кирлианами ещё в советский период, показали, что вид изображений меняется при изменении функционального состояния и самочувствия человека. Например, по виду кирлианограмм пальцев рук и ног человека оказалось возможным судить об общем уровне и характере физиологической активности его организма, по типу свечения — проводить классификацию его состояния, а по характеру распределения проекций свечения на акупунктурные каналы человека оказалось возможным оценивать состояние отдельных функциональных систем и органов человека, следить за влиянием на организм различных воздействий — лекарственных, лучевых биоэнергетических, вплоть до влияния на характер кирлианограммы состояний психомышечной релаксации человека под влиянием аутотренинга.

I. ФИЗИЧЕСКИЕ ПРОЦЕССЫ ФОРМИРОВАНИЯ ГРВ ИЗОБРАЖЕНИЙ

Газовый разряд возникает в системе, состоящей из объекта исследования, носителя изображения и электродов, формирующих электромагнитное поле (ЭМП). Первичным процессом является взаимодействие ЭМП со сканируемым объектом, в результате которого возникает эмиссия поверхности объекта заряженных частиц, участвующих в иницировании начальных фаз газового разряда при определенной напряженности ЭМП. Основным источником формирования изображения — это газовый разряд вблизи поверхности исследуемого объекта. Экспериментальные исследования показали, что можно выделить два основных типа разряда, связанных с формированием газоразрядных изображений: лавинный,

развивающийся в ограниченном диэлектриком узком зазоре, и скользящий по поверхности диэлектрика. Фотоны и заряженные частицы, возникающие в процессе разряда, формируют двумерную картину (Рис. 2) на носителе изображения. Газовый разряд, в свою очередь, может влиять на состояние объекта, вызывая вторичные эмиссионные, деструктивные и тепловые процессы. При исследовании микробиологических объектов было показано, что интенсивность большинства линий этой области зависит от состояния объекта. В оптической области спектра интенсивность линий существенно ниже, их положение и амплитуда зависят от типа объекта. Применение спектральных приборов различного типа позволило убедиться, что эти линии являются излучением объекта, а не переотражением. Эти эксперименты доказывают существенную роль в процессе ГРВ оптического излучения биологического объекта в видимой и ультрафиолетовой области спектра излучений сканируемого объекта.

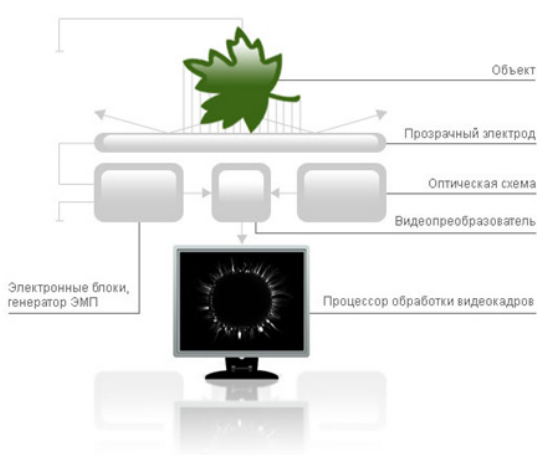


Рис. 1 – Принцип газоразрядной визуализации

Рисунок 1 иллюстрирует принцип метода ГРВ: Исследуемый объект помещается на поверхности диэлектрической пластины, на которую подаются высоковольтные импульсы от генератора, для чего на обратную сторону пластины нанесено прозрачное токопроводящее покрытие. При высокой напряженности поля в газовой среде пространства контакта объекта и пла-

стины развивается разряд в газовой фазе, носящий название «скользящий газовый разряд», параметры которого зависят от свойств исследуемого объекта. Свечение разряда фиксируется с помощью оптической системы и через интерфейс связи сохраняется в памяти компьютера в виде одиночных грв-грамм. Процессор обработки представляет собой программное обеспечение, которое позволяет вычислить комплекс параметров и на их основе делать определенные диагностические заключения.

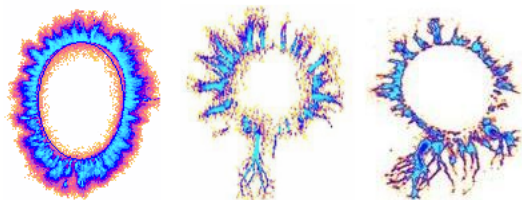


Рис. 2 - Примеры свечения большого пальца руки в поле высокой напряженности

II. ПАРАМЕТРЫ ГРВ-ГРАММ, ИСПОЛЬЗУЕМЫЕ ДЛЯ АНАЛИЗА

Существующие прикладные пакеты обработки изображений не могут быть использованы для обработки ГРВ-грамм в связи со спецификой возникающих задач, необходимостью учета диагностических гипотез и проведения обработки на уровне систем принятия решения. Поэтому была разработана программная среда обработки и анализа ГРВ-грамм, ориентированная на работу в различных предметных областях. Адаптация осуществляется за счет комбинации оптимальных для данной предметной области операций из библиотеки, выбора соответствующих процедур и (или) подбора оптимальных пороговых значений. В состав библиотеки включены следующие основные алгоритмы. Программная обработка ГРВ изображений включает несколько этапов: съемка ГРВ-грамм; фильтрация изображений (устранение шума); определение положения эллипсов с корректировкой центров и вычисление спектра интенсивности свечения в условных единицах от 0 до 255. Спектр изображения характеризует распределение количества пикселей для каждой единицы интенсивности. По спектру определяется порог уровня шума – удаляются все точки с интенсивностью меньше эмпирически выбранного уровня шума. После этого изображение разбивается на связанные фрагменты – участки изображения, у которых соседние пиксели примыкают друг к другу по вертикали или по горизонтали. Далее удаляются части изображения, у которых площадь фрагмента ни-

же 30 точек, после чего получается отфильтрованное изображение

III. ОБЛАСТЬ ПРИМЕНЕНИЯ МЕТОДА

Программно-аппаратные ГРВ биоэлектрографические комплексы нашли практическое применение в медицине.

- анализ психологического состояния личности;
- анализ психофизиологического состояния личности;
- анализ вегетативного статуса организма и отдельных функциональных систем;
- мониторинг реакций организма в процессе проводимой терапии;
- оценка вероятности наличия органических системных нарушений;
- оценка опасности аллергенов по параметрам ГРВ свечения образцов крови;

Одним из важных приложений метода ГРВ стала диагностики стрессовых состояний. Метод основан на регистрации свечения пальцев рук конкретного человека и сопоставления их с оптимальными параметрами, выявленными путем статистического анализа баз данных более 20000 съемок.

IV. ВЫВОДЫ

Исследования в области ГРВ биоэлектрографии активно продолжаются: регулярно защищаются диссертации, публикуются сборники научных трудов. Ежегодно, в июле, в Санкт-Петербурге проводится международная научная конференция «Наука, Информация, Сознание», на которую съезжаются врачи, ученые и специалисты из десятков стран мира. Последние годы эти конференции проходят под эгидой Международного Союза Медицинской и Прикладной Биоэлектрографии (IUMAB). Этот Союз объединяет исследователей из 62 стран, занимающихся практически внедрением методов биоэлектрографии, число которых с каждым годом расширяется

СПИСОК ЛИТЕРАТУРЫ

1. Коротков, К. Г. Принципы анализа ГРВ биоэлектрографии / К. Г. Коротков // Санкт-Петербург. – 2007. – 286 с.
2. Коротков, К. Г., Виллиамс Б., Виснески Л. А. Биофизические механизмы метода ГРВ биоэлектрографии / К. Г. Коротков, Л.А.Виснески // Санкт-Петербург. – 1998. – 344 с.
3. Стивен Смит. Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников. / Смит С.; пер. с англ. Линович. А.Ю., Витязев С.В. // Москва. – 2012. – 720 с.

АПЕРИОДИЧЕСКИЙ ЦИФРОВОЙ РЕГУЛЯТОР С ПРОГНОЗИРУЕМЫМ УРОВНЕМ СИГНАЛА УПРАВЛЕНИЯ

Олиферович Н. М., Гринюк Д. А., Оробей И. О., Сухорукова И. Г.

Кафедра автоматизации производственных процессов и электротехники, Белорусский государственный технологический университет

Минск, Республика Беларусь

E-mail: hope_rez@mail.ru, pryniuk@tut.by,

В работе представлены различные алгоритмы для расчета параметров настройки аperiodического цифрового (deadbeat) регулятора. Основное внимание при разработке алгоритмов расчета было уделено возможности построения данного регулятора для стабилизации технологических параметров в условиях ограничения на управляющее воздействие. Предложена методика расчета настроек deadbeat регулятора при условии формирования максимально возможных управляющих импульсов на начальной стадии отработки сигнала задания. Проанализирована зависимость максимально возможного управляющего воздействия от времени дискретизации при различных вариантах удлинения переходного процесса. Представлены результаты моделирования системы, которая была синтезирована на основании разработанных алгоритмов расчета.

ВВЕДЕНИЕ

В современных системах управления практически все алгоритмы реализуются в цифровом виде. При этом, зачастую, при стабилизации технологических параметров используются аналоговые методы синтеза. На сегодняшний день базовым алгоритмом стабилизации технологических параметров является пропорционально-дифференциально-интегральный (ПИД) закон регулирования. Однако данный алгоритм имеет ряд ограничений. Существует проблема возникновения интегрального насыщения при ограничении управляющего воздействия, проблема быстрогодействия, адаптация настроек при изменении динамических характеристик объекта.

При синтезе каскадных систем управления желательно иметь внутренний регулятор, который настроен на аperiodический переходной процесс типа deadbeat [1,2]. Это не только обеспечивает высокую скорость внутренней цепи (что является основной задачей внутренних регуляторов), но и облегчает задачу -за внешнего регулятора (наличие комплексных корней в объекте усложняет настройку внешнего регулятора). Алгоритмы deadbeat контроллера широко используются в управлении двигателем, в модуляции балансировки напряжения.

I. ОСНОВНАЯ ЧАСТЬ

Расчет многих цифровых регуляторов, в том числе deadbeat [1,2], гораздо более алгоритмизирован. Данный регулятор характеризуется высоким быстродействием. Однако это достигается путем формирования высокого значения управляющего воздействия, что сложно реализовать в прикладных условиях. В то же время можно добиться соответствия контроллера технологическим нормам за счет увеличения времени шаговых ответов.

Следует отметить, что существуют и другие подходы к проектированию deadbeat регуляторов [3-5]. Однако в прикладных решениях находит большее применение подход из работы [1]. Синтез производился на основании функции z -преобразования элементов систем управления. Общая структура системы управления показана на рисунке 1.

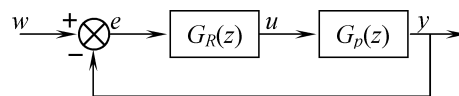


Рис. 1 – Структура deadbeat регулятора: w – сигнал задания; e – сигнал ошибки; u – управляющее воздействие; y – выход системы

Передаточная функция $G_R(z)$ регулятора и объекта может быть записана следующим образом [1]

$$G_R(z) = \frac{Q(z)}{1 - P(z)}; G_P(z) = \frac{A(z)}{B(z)}$$

где $Q(z)$, $P(z)$ – многочлены передаточной функции регулятора; $A(z)$, $B(z)$ – многочлены передаточной функции объекта. Коэффициенты многочленов $Q(z)$, $P(z)$ для классического deadbeat регулятора могут быть найдены с использованием уравнений [1]:

$$\begin{cases} q_0 = \frac{1}{b_1 + b_2 + \dots + b_m}, \\ q_i = q_0 a_i, \\ p_i = q_0 b_i, \end{cases} \quad (1)$$

где a_i , b_i коэффициенты многочленов функции z -передачи непрерывного объекта $A(z)$, $B(z)$; m порядок передаточной функции объекта. Коэффициенты многочлена $Q(z)$ можно выразить так:

$$\begin{cases} q_0 = u(0), \\ q_0 = u(i) - u(i-1), i = 1, \dots, m. \end{cases} \quad (2)$$

Таким образом, начальное значение управляемой переменной $u(0)$ зависит только от суммы (1) коэффициентов функции z -передачи объекта b_i . Основным недостатком deadbeat регулятора является то, что время выборки t_s и управляемая переменная $u(0)$ обратно пропорциональны. Значение t_s обусловлено величиной показателя u_{max} , которая зависит от различных факторов. Для решения этой проблемы предлагаются различные варианты, такие как изменение алгоритма расчета. Авторы вводят дополнительный полином для ограниченного вывода [6]. Но могут быть и другие решения. Возможны системы управления с переменным временем выборки t_s . Для синтеза системы, выражения, полученные в работе [1] для выбора первичного импульса, были расширены на большее количество тактов.

II. МОДЕЛИРОВАНИЕ И ОЦЕНКА РОБАСТНОСТИ

Для оценки работы синтезированных регуляторов был использован объект с передаточной функцией

$$W(s) = \frac{(2s + 1)}{(10s + 1)(7s + 1)(3s + 1)} \exp(-4s) \quad (3)$$

Некоторые из результатов моделирования синтезируемой системы представлен на рис. 2–4, где отображено влияние количества одинаковых управляющих воздействия первых тактов N на максимальное управляющее воздействие при переходном процессе.

Оценка устойчивости произведена по диаграммам Боде с определением запасов по амплитуде A и фазе (рис. 5–6).

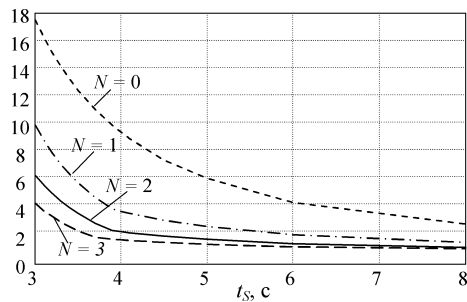


Рис. 2 – Зависимость первого максимального управляющего воздействия u_{max} от $t_s=3,7$ с

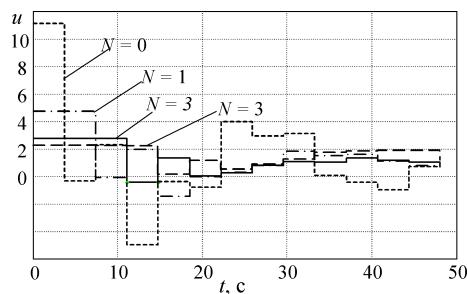


Рис. 3 – Переходной процесс сигнала управления u при $t_s=3,7$ с

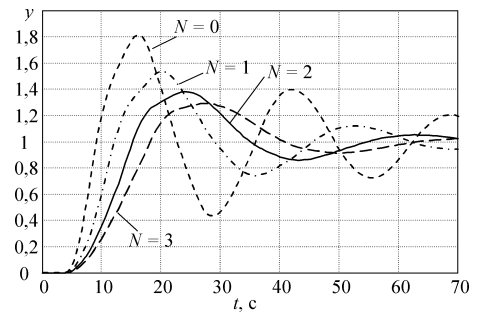


Рис. 4 – Переходной процесс выходного сигнала y при $t_s=3,7$ с

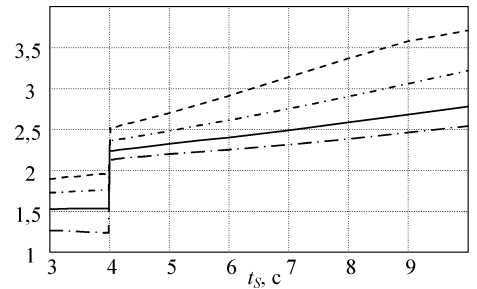


Рис. 5 – Зависимость запаса по амплитуде A от t_s при разном N

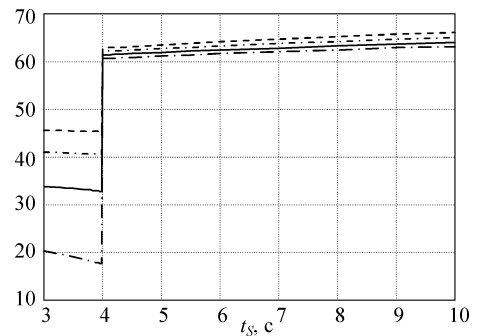


Рис. 6 – Зависимость запаса по фазе φ от t_s при разном N

Следует отметить, что при t_s меньше величины запаздывания наблюдается неоднозначная зависимость запасов по фазе и амплитуде от t_s . При некоторых значения можно было наблюдать неустойчивый режим работы замкнутой системы.

1. Изерман Р. Цифровые системы управления: пер. с англ. М.: Мир, 1984. 541 с.
2. Plsek S., Matus R. Application of Adaptive Deadbeat Controller in Drying Process // Procedia Engineering. 2015. Vol. 100. P. 756–764.
3. Цифровые алгоритмы для управления технологическими процессами/ Д. А. Гринюк [и др.] // Информационные технологии и системы–2013 (ИТС–2013): материалы междунар. науч. конф., Минск, 23 окт. 2013 г. Белорус. гос. ун-т информатики и радиоэлектроники, Минск, 2013. С. 38–39.
4. Balasevicius L., Dervinis G. Design of the Deadbeat Controller with Limited Output // Electronics and electrical engineering. 2011. No. 4 (110). P. 93–96.
5. Kučera V. Deadbeat Control, Pole Placement, and LQ Regulation // Kybernetika. 1999. Vol. 35, no. 6. P. 681–692. L., Baranauskas V., Dervinis G., Derviniene A. Deadbeat Controller with Two Additional Steps. Elektronika ir Elektrotechnika, 2017, vol. 23, no. 5, pp. 11–14.

ПРОБЛЕМЫ БЕЗОПАСНОСТИ SIP-ТЕЛЕФОНИИ

Полудворянин С. М., Нестеренков С. Н.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: siarheipoludvaranin@gmail.com, nsn@bsuir.by

В последние годы протокол SIP получил широкое распространение. Среди основные преимуществ SIP-телефонии перед традиционной телефонией выделяют снижение затрат, масштабируемость, гибкость и надежность. Однако миграция на SIP-телефонию создает новые проблемы для организаций, которые не существовали при использовании телефонной сети общего пользования. В частности, необходима задуматься о безопасности SIP-телефонии.

ВВЕДЕНИЕ

Время аналоговой телефонии (PSTN) подходит к концу. Привычные телефонные линии переводят на современную гибкую технологию – SIP-телефонии (Session Initiation Protocol). SIP является протоколом управления прикладного уровня для создания, изменения и завершения сеансов связи с одним или большим количеством участников. В понятие сеанса входят мультимедиа конференции, обучение на расстоянии, Internet-телефония и подобные приложения[1]. В ближайшем будущем неизбежен полный переход аналоговых телефонных линий на IP-телефонию, а корпоративных АТС на протокол SIP. Согласно аналитическому отчету компании IHS количество SIP-транков в мире за первое полугодие 2017 года выросло 26% по сравнению с прошлым годом и достигло 30 миллионов штук[2]. Однако вместе с распространением технологии мы все чаще слышим сообщения о взломах и кибер-атаках на телефонные сервисы. В этой статье мы рассмотрим несколько угроз, возникающих при использовании протокола SIP, и способы эти угрозы реализовать.

I. НЕАВТОРИЗИРОВАННЫЙ ДОСТУП К ДАННЫМ

Открытые текстовые протоколы предоставляют всю информацию любому, кто способен перехватывать сетевой трафик. Анализируя полученные данные можно получить доступ для конфиденциальной информации. Возьмем к примеру процесс аутентификации. В sip-протоколе пароли кодируются с помощью алгоритма MD5[1], который уже давно считается небезопасным. Множество телекоммуникационного оборудования, а также программные решения, поставляются с известными паролями по умолчанию. Если эти пароли остаются без изменений, злоумышленники могут легко получить их. Полученные данные аутентификации могут быть использованы для подмены легальной регистрации SIP-аккаунта на серверах VoIP провайдера (Registration hijacking). Полученный SIP-аккаунт может быть использован для соверше-

ния огромного количества звонков на специальные дорогие номера, а также рассылки аудио спама. Суммы, которые придется заплатить владельцу скомпрометированного аккаунта, могут быть весьма и весьма значительными. В некоторых случаях аутентификация может осуществляться на основе известной пары IP-адреса и порта. В случае если у клиента-жертвы нет выделенного IP-адреса, злоумышленник, находящийся с ним в одной локальной сети, получает возможность осуществлять звонки за счет жертвы. Такая ситуация может возникнуть, если провайдер осуществляет доступ к сети Интернет через NAT (от англ. Network Address Translation – «преобразование сетевых адресов»), за которым помимо законного абонента сервиса IP-телефонии могут находиться другие пользователи. В качестве примера можно привести бизнес-центр, где большое количество фирм проходят через NAT местного провайдера.

II. ПЕРЕХВАТ И ЦЕЛОСТНОСТЬ ДАННЫХ

Для прослушивания аналоговой телефонной линии злоумышленнику потребуется получить физический доступ к линии или коммуникационному оборудованию. В IP-телефонии возможность прослушивания разговоров резко увеличилась благодаря большому количеству сетевых узлов между субъектами разговора. Скомпрометировав хотя бы одну сетевую ноду, атакующий получит доступ к медиа-трафику. На данный момент с сети Интернет достаточно бесплатного программного обеспечения для преобразования VoIP-трафика в аудиофайлы и их прослушивания. Полученную таким образом информацию очевидно можно использовать с целью шантажа или шпионажа или применить в социальной инженерии. Для SIP-протокола характерно отсутствие взаимной аутентификации в большинстве случаев обмена сообщениями между сервером и клиентом. Без взаимной аутентификации злоумышленник, осуществляющий атаку типа «человек посередине»[3], может подменять исходящие/входящие пакеты клиента-жертвы, нарушая целостность передаваемой информации. Получив возможность для подме-

ны данных, злоумышленник может перенаправлять звонки клиента на произвольный номер, в том числе на несуществующий (приводя к DoS). Перехватывая INVITE-пакеты, злоумышленник подменяет ответ пакетом с кодом 301 Moved Permanently, перенаправляя звонок. Злоумышленник, зная или подбирая некоторые атрибуты звонка, способен разрывать активные сессии, отправляя заранее сформулированные вредоносные пакеты, содержащие запрос на завершение соединения(SIP BYE-request)[4].

III. ОГРАНИЧЕНИЕ ДОСТУПНОСТИ

Для сервисов IP-телефонии крайне важно быть всегда доступными для клиентов и предоставлять хорошее качество связи. Именно поэтому они так подвержены атакам типа «отказ в обслуживании» (Denial of Service, DoS). Атака DoS может быстро снизить уровень QoS до неприемлемого уровня. Эта атака нацелена на превышение предельной нагрузки на систему большим количеством коротких звонков или информационного мусора. Если для атаки создается сеть из большого количества «зомби-компьютеров», то такую атаку называют DDoS-атакой (от англ. Distributed Denial of Service, распределённая атака типа «отказ в обслуживании»). DoS-атаки могут иметь различные формы и мотивацию. Наиболее простой способ атаки на SIP сервер это инициация большого количества сессий с различными идентификаторами с целью исчерпания оперативной памяти сервера. Различные типы SIP-запросов (REGISTER, INVITE или OPTIONS) могут быть использованы для такого вида атаки[5]. Другой сценарий, который может использовать злоумышленник, изображен на Рис. 1. В этом случае злоумышленник пытается «вести себя» как законный пользователь. Скорее всего, злоумышленник будет пытаться реализовать различные сценарии с отправкой множества INVITE запросов. Это приводит к отказу в обслуживании сервера SIP провайдера или АТС жертвы. Без постоянного отслеживания признаков подобных атак и применения пассивных средств защиты, это приводит к тому, что серверы IP-телефонии не справляются с возросшей нагрузкой и не в состоянии обслуживать подключенных абонентов. DDoS атака может применяться и на медиа-сервера. Она осуществляется путем отправки большого количества RTP-пакетов. Если цель не способна обработать такое количество пакетов, то качество связи ухудшается, либо сервер перестает работать. Поскольку большинство реализаций RTP базируется на UDP, такой трафик легко генерировать. Выход из строя одного медиа-сервера мо-

жет нарушить работу передачи голосового трафика всей сети.

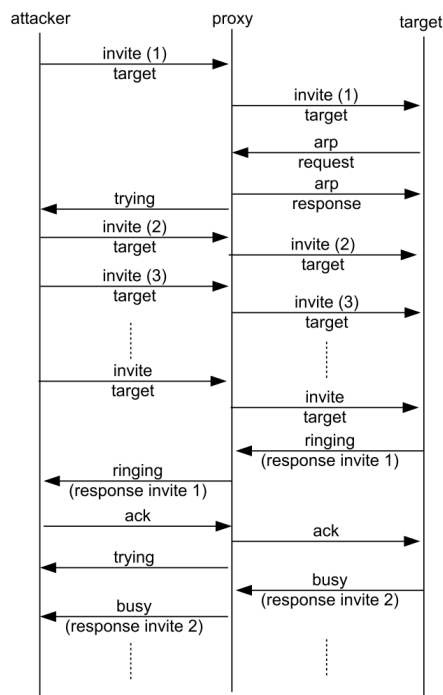


Рис. 1 – DDoS INVITE-запросами

IV. ЗАКЛЮЧЕНИЕ

В этой статье мы рассмотрели основные проблемы безопасности IP-телефонии и виды угроз, с которыми можно встретиться. Клиенты и провайдеры SIP-телефонии должны серьезно относиться к требованиям безопасности и конфиденциальности своих голосовых звонков. Кроме того, они должны учитывать последствие в случае взлома.

1. Гольдштейн, Б. С. IP-Телефония / Б. С. Гольдштейн, А. В. Пинчук, А. Л. Суховицкий. – СПб.: БХВ, 2014. – 336 с.
2. SIP Trunks to Top 53 Million in 2021 [Electronic resource] / D. Myers – IHS Markit, 2017. – Mode of access: <https://technology.ihs.com/596390/sip-trunks-to-top-53-million-in-2021>. – Date of access: 14.08.2018.
3. Нестеренков, С. Н. Защита информации в приложениях масштаба предприятия / С. Н. Нестеренков, Б. В. Никульшин // Технические средства защиты информации : тез. докл. VIII Белорус.-рос. науч.-техн. конф., Браслав, 24-28 мая 2010 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. М. Лыньков [и др.]. – Минск, 2010. – С. 60–61.
4. Lohiya, K. End to End Encryption Architecture for Voice over Internet Protocol /K. Lohiya, N. Shekar, S. R. Devane //International Journal of Computer Applications – 2012. – Vol. 41. – P. 31–34.
5. Wang, Y. A Practical Method for SIP-DoS Attack Effect Evaluation /Y. Wang, Y. Yang //Journal of Information & Computational Science – 2012. – Vol. 9. – P. 333–345.

ПРОЦЕССЫ ИСТЕЧЕНИЯ И СВОДООБРАЗОВАНИЯ В БУНКЕРАХ

Пригара В. Н.

Кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: prigara@bsuir.by

Бесперебойное и равномерное истечение груза из емкостей в значительной степени зависит от следующих факторов: режима функционирования бункера, конфигурации и конструкции корпуса и выгрузной воронки, месторасположения и размера выпускного отверстия и физико-механических свойств груза.

Ключевые слова: гравитационный выпуск сыпучего груза; слеживаемость; сводообразование

Механизация погрузочно-разгрузочных работ – один из важнейших резервов повышения экономической эффективности технологических операций с сыпучими грузами. Многообразие насыпных грузов, различающихся по свойствам и назначению, обуславливает применение различных бункеров, бункерных устройств, загрузочных приспособлений, стабилизаторов истечения, питателей и т. д. Все они относятся к системе транспортирующих механизмов для комплексной механизации грузопотока. На гравитацион-

ное истечение сыпучего груза из отверстия емкости влияют многие факторы, которые могут быть сгруппированы так: режимные (технологические) паузы в загрузке, выпуске, физико-механические свойства грузов и параметры элементов емкости. На стабильность истечения грузов, а следовательно, и процесс сводообразования, оказывают влияние повышенное содержание влаги и уплотнение груза над выпускным отверстием. [1]



Рис. 1 – Классификация факторов, влияющих на гравитационный выпуск сыпучего груза из емкостей

Грузы, теряющие подвижность частиц при длительном хранении, называются слеживающимися. Хранение в течение нескольких суток связных грузов вызывает увеличение начального сопротивления сдвигу в 5-8 раз. При отрицательной температуре окружающей среды хранение влажных грузов сопровождается смерзанием. Смерзание усложняет погрузку и выгрузку грузов.[2]

При истечении насыпного груза из бункера через разгрузочное отверстие образуется свод из подвижных частиц, воспринимающий давления вышележащих слоев. В первой зоне, лежащей над сводом, частицы движутся в замкнутом слое и соприкасаются друг с другом. Из этого слоя частицы непрерывно переходят в область самого свода. Под поверхностью свода начинается зона свободного падения частиц под действием собственной силы тяжести. В этой второй зоне уже нет связанных друг с другом слоев частиц, они не соприкасаются, при падении расстояния между ними увеличиваются.

Слеживаемость (англ. - Slumping) (или липкость, англ. - stickiness, tackiness) - это способность отдельных частиц груза сцепляться, прилипать к стенкам подвижного состава, бункеров и друг к другу и образовывать достаточно прочную монолитную массу, т. е. терять сыпучесть или взаимную подвижность частиц при длительном хранении.

Явление слеживаемости следует рассматривать как проявление сцепления частиц насыпных грузов. Динамические нагрузки ускоряют процесс слеживаемости. При этом выпуск грузов значительно затрудняется. Использование для побуждения истечения ударных нагрузок приводит к образованию пустот (устойчивых статических сводов) над выгрузным отверстием. Устойчивость пустот зависит от сил аутогезии частиц и площади поперечного сечения выпускного отверстия.[1]

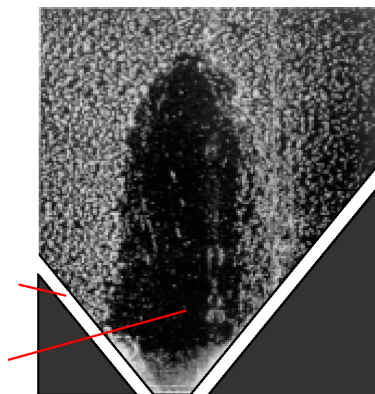


Рис. 2 – Слежавшийся груз после длительного хранения

Процесс истечения сыпучих грузов из бункеров нарушается, в первую очередь, вследствие возникновения явлений сводообразования. Особенно сильно этому явлению подвержены связные грузы.

На основании исследований известно три типичных формы истечения грузов различной степени связности из бункерно-силосных емкостей:

- свободное истечение сыпучего груза малой связности из выгрузного отверстия диаметром d с образованием потока диаметром d ($d/d > 1$), обрушением откосов в образующуюся воронку и наличием незначительных остатков на днище емкости;
- истечение сыпучего груза повышенной связности с пульсацией и образованием динамических сводов в зоне выгрузки;
- отсутствие гравитационного выпуска груза высокой связности из-за статических сводов при $d < d$. Разгрузка емкости производится с применением специальных выгрузных устройств. [1]

СПИСОК ЛИТЕРАТУРЫ

1. И.В. Горюшинский, И.И. Кононов Емкости для сыпучих грузов в транспортно-грузовых системах, Учебное пособие, Самара, 2003
2. Р.Л. Зеленков, Г.П. Гриневич, В.С. Исаев Бункерные устройства, Москва, 1977

ОБЗОР И СРАВНЕНИЕ СУЩЕСТВУЮЩИХ МЕТОДОВ ОЧИСТКИ ПОВЕРХНОСТЕЙ

Пригара В. Н.

Кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: prigara@bsuir.by

Под воздействием множества факторов сыпучие материалы, которые обычно хранятся в накопительных бункерах, могут образовывать своды и зависание вещества. При этом поток выдачи сырья либо меньше номинального, либо вообще прекращается, что является сдерживающим фактором непрерывного технологического процесса.

Ключевые слова: ручная очистка; механические вибрации; газо-импульсная очистка; магнитное поле

В настоящее время существует несколько технологий, отличающихся по своей физической природе, которые используются в целях очистки рабочих поверхностей от налипших материалов. Отличия обусловлены, прежде всего, используемым рабочим телом. Наиболее применяемыми являются следующие рабочие тела:

- человеческий труд;
- механические вибрации;
- газ;
- магнитное поле.

У каждой из технологий, использующих данные рабочие тела, имеются свои преимущества и недостатки, совокупность которых на практике формирует конкретную область применения метода [1].

1. Ручная очистка

Под ручной очисткой понимается очистка силами рабочих, обслуживающих данный бункер. Для того, чтобы выходной поток сырья был постоянным, необходим непрерывный контроль за ходом технологического процесса. А так как данный метод обладает крайне низкой перспективой автоматизации, то существует необходимость выделения нескольких работников за наблюдением целого парка бункеров на предмет сводообразования или зависания сырья.

Как правило, очистка производится с внешней стороны бункера при помощи кувалд или длинных шестов. Соответственно деформации, вызываемые в бункере, практически не управляемы по своей величине и месту приложения, и вполне могут носить неупругий характер.

2. Механические вибрации

Источниками механических вибраций могут служить разные устройства, однако на практике наиболее часто употребляются промышленные вибраторы. Конструкция вибратора обеспечивает виброустойчивость статорной обмотки и механическую прочность вала ротора, подшипникового узла и корпусных деталей. Статор электродвигателя встроен в алюминиевый или чугунный литой корпус с коробкой выводов и усиленными элементами крепления к вибромеханизму.

Вал ротора опирается на подшипники. При вращении ротора электродвигателя возникают круговые колебания вибратора, присоединенного к нему механизма.

По применяемой энергии различают следующие виды вибраторов:

- электромеханические, в которых вибрация происходит за счет вращения дебалансов, установленных на валу электродвигателя;
- электромагнитные, в которых колебания совершаются с помощью электромагнитов постоянного и переменного тока;
- пневматические, работающие при помощи сжатого воздуха.

Для обрушения сводов с помощью вибраторов необходимо знать величину распространения колебаний, которая зависит от конструкции емкости и физико-механических свойств его содержимого. Интенсивность колебаний в материале убывает относительно быстро. Исследования показывают, что прямолинейные колебания от электромагнитных вибраторов распространяются в песке на расстояние до 1,5 м [2].

Из основных преимуществ данного метода можно отметить простоту монтажа и долгий срок службы без особых затрат на техническое обслуживание. Из недостатков: высокую стоимость, низкую эффективность, сложность в изменении параметров вибрации, а также работу в области неупругих деформаций бункера.

3. Газо-импульсная очистка

Все газо-импульсные технологии основаны на кратковременном воздействии мощных газовых струй, которые создаются при помощи специальных газо-импульсных генераторов.

Технология газо-импульсного обрушения предназначена для устранения зависаний и налипаний материалов на стенках рабочих емкостей, независимо от их назначения, размеров и конструкции. В основу положено ударно-волновое воздействие воздуха на материал в емкости, которое осуществляется с помощью газо-импульсного генератора. Генератор заполняется сжатым воздухом от обычной воздушной систе-

мы в течение 10 - 20 секунд, а затем выбрасывает накопленный воздух за доли секунды, что позволяет увеличить секундный расход в сотни раз и организовать мощное импульсное воздействие на обрабатываемый материал.

4. Очистка посредством магнитного поля

Наиболее эффективной технологией для решения проблем налипания, зависания и сводообразования в настоящее время является магнитно-импульсная установка, основанная на использовании силового воздействия импульсного магнитного поля на электропроводные материалы.

Силовой блок генерирует мощный импульс тока в обмотку индуктора. Магнитное поле индуктора, созданное этим током, индуцирует импульс тока в плите, установленной вблизи индуктора. В результате взаимодействия импульсных токов, протекающего по обмотке индуктора и наведенного в плите, плита оказывает импульсное механическое воздействие на очищаемую поверхность, что приводит к возникновению локальной упругой деформации в очищаемой поверхности, а в толще налипшего материала – к возникновению напряжений сдвига. Совместное действие этих процессов нарушает целостность слоя налипшего материала, разрушает адгезию материала к очищаемой поверхности и приводит к ее очистке. Сила механического воздействия и количество импульсов регулируются и выбираются достаточными для гарантированного обрушения налипших материалов.

При очистке поверхности (стенки бункера) последовательно на каждый задействованный канал установки подается серия импульсов, количество импульсов в серии и интервал между ними регулируются. После последовательной подачи серий импульсов на все задействованные каналы установка переходит в режим ожидания. Следующая подача импульсов в исполнительные механизмы происходит согласно установленному режиму работы (либо от ручного сигнала, либо от сигнала датчика, либо по установленной программе) [3].

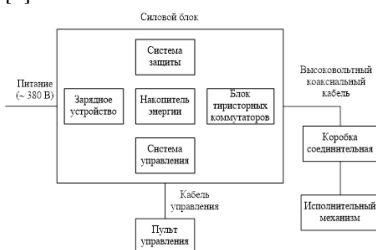


Рис. 1 – Примерная структурная схема магнитно-импульсной установки

Практика эксплуатации магнитно-импульсных установок на различных предприятиях свидетельствует о более высокой эффективности магнитно-импульсного способа сводообрушения и очистки по сравнению с вибраторами

(виброднищами) и пневмообрушением. Технологии очистки на основе магнитно-импульсной установки являются энергосберегающими.

5. Сравнение и обоснование выбора типа системы обрушения

Сведем все полученные данные и проанализируем полученные результаты. Наиболее оптимальным выбором из всех существующих методов очистки поверхностей от налипших материалов, является метод магнитно-импульсной очистки. Основные преимущества данного метода таковы:

- Более высокая эффективность очистки, реализуемая магнитно-импульсными установками, по сравнению с другими системами (вибраторы, пневмообрушение), связанная с возможностью согласования амплитудно-частотных характеристик воздействующего импульса с физикотопологическими параметрами бункера и сыпучего материала и, как результат, с достижением гарантированного обрушения налипшего материала с минимальными энергетическими затратами.
- Уменьшение эксплуатационных затрат по содержанию систем сводообрушения. В отличие от электромеханических вибраторов магнитно-импульсные установки являются энергосберегающими.
- Повышение безопасности труда за счет уменьшения, а в некоторых случаях исключения необходимости применения ручного труда для очистки бункеров и других объектов.
- Обеспечение целостности стенок бункеров при их очистке.
- Надежность и долговечность магнитно-импульсных систем за счет отсутствия в исполнительных механизмах соударяющихся, вращающихся и трущихся частей, применения оригинальных схемных решений, присутствия целого ряда защит от нештатных режимов.
- Магнитно-импульсные установки не оказывают вредного влияния на чувствительные элементы контрольно-измерительной аппаратуры.

Таким образом, из вышеприведенных типов очистки поверхностей от налипших материалов выбираем магнитно-импульсный тип системы обрушения, поскольку по основным характеристикам именно этот тип систем является наилучшим для поставленной задачи.

1. Пригара В.Н. Исследование режимов очистки бункеров от солеотложения, ITS 2012, 2012
2. Промтов М.А. Машины и аппараты с импульсными энергетическими воздействиями на обрабатываемые вещества, «Издательство Машиностроение -1», 2004
3. Пригара В.Н. Магнитно-импульсная установка, 48-я научная конференция аспирантов, магистрантов и студентов, 2012

УСТРАНЕНИЕ КОНФЛИКТОВ МАРШРУТОВ ПРИ ОБЛЕТЕ ЦЕЛЕЙ ГРУППОЙ БПЛА

Симаньков В. И.

Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: demeka-v@mail.ru

При построении маршрутов для облета целей группой БПЛА иногда возникнут пересечения, что может привести к столкновениям. Получить такое решение, при котором сумма длин маршрутов близка к минимально возможной и при этом без пересечений это сложная математическая задача. Для ее упрощения предлагается решить задачу без ограничений, а затем разнести пересекающиеся маршруты на разные высоты, число которых должно быть минимальным.

ВВЕДЕНИЕ

Обобщением задачи построения маршрутов облета целей для группы БПЛА является задача нескольких коммивояжеров с ограничением на пересечения маршрутов. Решения задачи нескольких коммивояжеров без ограничений представлены в [1-3]. Предлагается использовать существующие решения для построения маршрутов, а для устранения их пересечений использовать следующий алгоритм.

1. ОПИСАНИЕ АЛГОРИТМА

Рассмотрим алгоритм пошагово на примере. Пусть на плоскости произвольно заданы начальные координаты m БПЛА с назначенными маршрутами, состоящими из K целей каждый. Исходные условия изображены на рис.1.

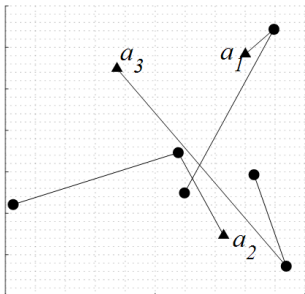


Рис. 1 – Исходные маршруты для $m = 3, K = 2$

Данные о пересечении маршрутов могут быть представлены в виде графа пересечений $E[m][m] \in [0,1]$, то есть, если элемент $E_{ij} = 1$, значит маршруты i -го и j -го БПЛА пересекаются. Здесь вершинами графа являются БПЛА, а ребра графа – соединяют БПЛА, маршруты которых пересекаются. Для разнесения пересекающихся маршрутов на разные высоты можно применить алгоритм расцветки вершин графов минимальным количеством цветов. Решением этого алгоритма является назначение каждой вершине цветов таким образом, что любые две смежные вершины имеют разный цвет. В случае с маршрутами БПЛА цвета означают уровни высот. На рис. 2 изображен граф пересечений, соответствующий исходным условиям на рис. 1.

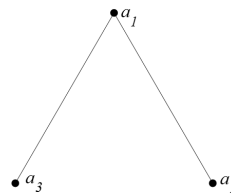


Рис. 2 – Соответствующий примеру $m = 3, K = 2$ граф пересечений

Задача минимальной расцветки графа сводится к целочисленному линейному программированию (ЦЛП) [4]. Номер цвета представлен целочисленными переменными v_i , ассоциированными с вершинами (БПЛА).

$$\begin{cases} v_i \in [1, m] \\ i \in [1, m] \end{cases}, v_i \in \mathbb{Z}, \\ z = \max v_i \rightarrow \min$$

где m – количество БПЛА
 z – целевая функция.

Для удобства введена еще одна целочисленная переменная v_0 , которая связана с остальными системой неравенств

$$\begin{cases} v_i \leq v_0 \\ i \in [1, m] \end{cases}, v_0 \in \mathbb{Z}, \quad (1) \\ z = v_0 \rightarrow \min$$

В задаче о правильной раскраске есть ограничение: соседние вершины должны иметь разные цвета. Это значит: для каждого ребра $e_{ik} \in E$ переменные v_i, v_k должны отличаться хотя бы на единицу и не более чем на $m - 1$: Чтобы записать это ограничение как систему линейных неравенств, автор [4] предлагает следующую систему неравенств

$$\begin{cases} v_i - v_k - ne_{ik} \leq -1 \\ v_k - v_i + ne_{ik} \leq m - 1 \end{cases}, \begin{cases} e_{ik} \in 0 \cup 1 \\ \forall e_{ik} \in E \end{cases} \quad (2)$$

где n – количество ребер.

Сформируем задачу ЦЛП. Необходимо минимизировать функцию z (1), которая зависит от $n+m+1$ переменных $v_0, v_1 \dots v_m, \forall e_{ik}$. На каждую v_i наложено ограничение z (1) всего m ограничений. Для каждой переменной e_{ik} и соответствующих ей v_i, v_k должны выполняться по 2

ограничения (1) – всего $2n$ ограничений. В общем виде данная задача записывается как

$$\min_x z = cx \begin{cases} x \in [1, m], x \in Z \\ A \cdot x \leq b \\ lb \leq x \leq ub \end{cases}$$

Для примера, изображенного на рис. 1, матрицы имеют вид

$$A = \begin{pmatrix} 0 & 1 & -1 & 0 & -3 & 0 \\ 0 & -1 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & -1 & 0 & -3 \\ 0 & -1 & 0 & 1 & 0 & 3 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

$$b = \begin{pmatrix} -1 \\ 2 \\ -1 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$lb = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, ub = \begin{pmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 1 \\ 1 \end{pmatrix}, m = 3, n = 2$$

Эта задача ЦЛП решается одним из методов либо использовать существующий решатель. Например, *intlinprog* из САПР MATLAB. Решение следующее

$$x_{min} = (2 \ 1 \ 2 \ 2 \ 0 \ 0)^T \Rightarrow v_0 = 2, v_1 = 1, v_2 = 2, v_3 = 2.$$

Это значит, что общее число слоев два, маршрут БПЛА a_1 распределен на первый слой, a_2, a_3 – на второй. Решение также представлено на рис. 2.

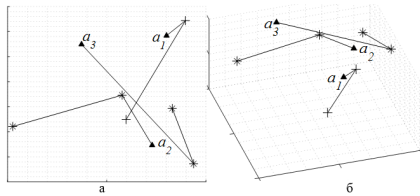


Рис. 3 – Разрешение конфликтов маршрутов разнесением по высотам с помощью алгоритма расцветки графов а – двумерная проекция, б – трехмерная проекция

Алгоритм подходит и для более размерных задач, например $m = 30, K = 2$. Исходные условия

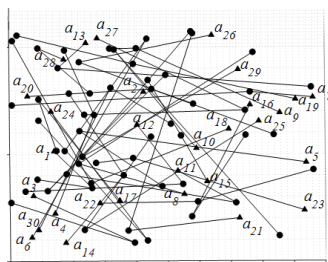


Рис. 4 – Исходные условия $m = 30, K = 2$

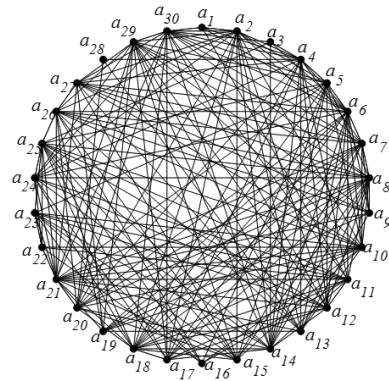


Рис. 5 – Граф $m = 30, K = 2$

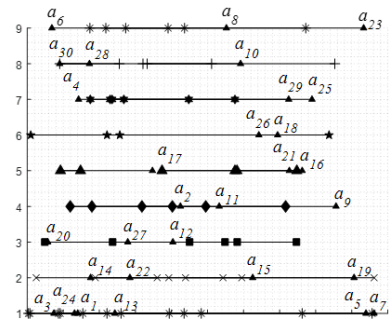


Рис. 6 – Решение $m = 30, K = 2$

II. ЗАКЛЮЧЕНИЕ

Предложенный алгоритм полезен тем, что нет ограничения на пересечения маршрутов. Поэтому можно их строить по критерию минимальной длины или минимального времени облета. Например, решить задачу нескольких путешествующих продавцов, решение которой обеспечивает кратчайший путь, но в общем случае содержит взаимопересечения маршрутов. А затем с помощью метода расцветки графов разнести пересекающиеся маршруты на минимальное количество слоев и тем самым получить бесконфликтное решение.

III. СПИСОК ЛИТЕРАТУРЫ

1. Yousefikhoshbakht, Majid Didehvar, Farzad Rahmati, Farhad. Modification of the Ant Colony Optimization for Solving the Multiple Traveling Salesman Problem. Romanian Journal of Information Science and Technology. №16 (2013). p 65-80.
2. Rainer Burkard Assignment Problems Mauro Dell'Amico Silvano Martello Society for Industrial and Applied Mathematics Philadelphia, PA, USA ©2009.
3. A New Algorithm for Solving Linear Bottleneck Assignment Problem Pramendra Singh Pundir1, Sandeep Kumar Porwal2 and Brijesh P. Singh3 Journal of Institute of Science and Technology, 2015, 20(2): 101-102,
4. Иглин, С. П. Математические расчеты на базе MATLAB // СПб.: БХВ-Петербург, 2005 – 640 с.

РАЗРАБОТКА УСТРОЙСТВА ДЛЯ АВТОМАТИЧЕСКОЙ НАМОТКИ ТРАНСФОРМАТОРА

Ситкевич Т. А.

Кафедра электротехники и электроники, Гродненский государственный университет им. Я. Купалы
Гродно, Республика Беларусь
E-mail: sittan@tut.by

Рассмотрены структурная схема устройства для автоматической намотки трансформатора с использованием датчиков Холла и принцип его работы. Для автоматизации процесса намотки трансформатора достаточно использования микроконтроллера ATmega328.

Целью данной работы является создание устройства для автоматической намотки трансформаторов с использованием датчика Холла, имеющего низкую себестоимость, требующего минимального времени при настройке и имеющего высокий уровень автоматизации. Объектом исследования являются принципы управления электродвигателями.

Главные задачи при создании устройства направлены на оптимизацию всех компонентов устройства для их компактной работы; программирование контроллера; установку автоматического контроля над скоростью намотки; создание 3d моделей устройства; сборку и компоновку устройства.

Структурная схема устройства приведена на рисунке 1.

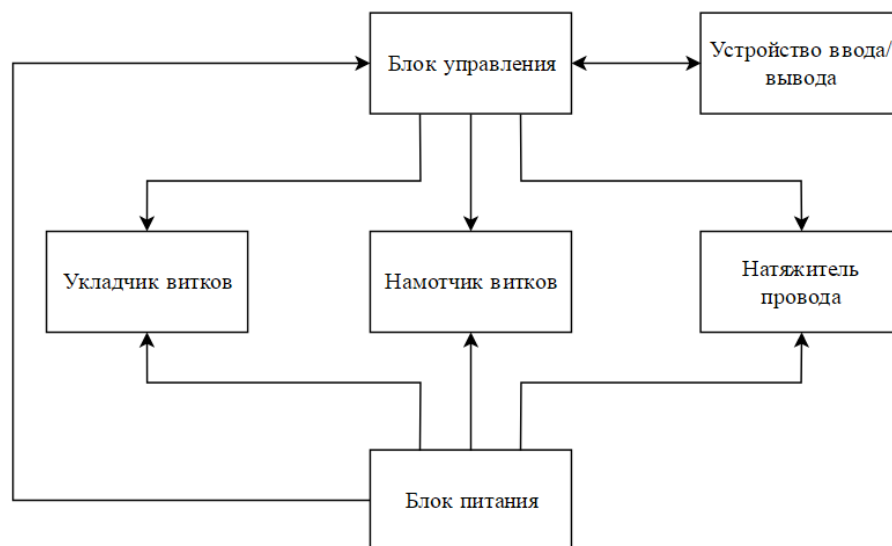


Рис. 1 – Структурная схема устройства автоматической намотки

Аппаратная часть проекта представляет собой контроллер, управляющий процессом намотки с помощью драйвера управления, основываясь на данных со счётчика витков на базе датчика Холла, устройства ввода данных и конечных выключателей. Программный код контроллера разработан в среде разработки Arduino IDE v1.6.9 с использованием встроенного компилятора и языка программирования Arduino [1].

При включении устройства происходит установка укладчика провода в начало координат, после ввода данных начинается процесс намотки. Намотка продолжается до тех пор, пока количество намотанных витков не достигло требуемого. Кроме того, при достижении укладчика конечных выключателей, изменяется направ-

ление укладки провода. При включении режима послойной намотки ожидается ввод команды на продолжение работы устройства после намотки одного слоя. В качестве счётчика витков используется цифровой датчик Холла А3144. Диапазон напряжений питания составляет 4.5 – 24 В. Сигнальный вывод датчика представляет собой открытый коллектор. При попадании чувствительного элемента датчика в магнитное поле сигнальный вывод соединяется с общим проводом через n-p-n-переход, и датчик отправляет соответствующий сигнал на один из входов контроллера [2].

В качестве привода укладчика провода используется шаговый двигатель ПБМГ 200-265Ф униполярного типа. Двигатель имеет две обмот-

ки с отводами от средних точек и, соответственно, шесть выводов. Для одного полного оборота вала двигателя требуется совершить 200 шагов, при этом за один шаг вал двигателя совершает поворот на 1,8 градуса. При сопротивлении обмоток в 70 Ом потребляемый ток двигателя без нагрузки на валу составляет 0,17 А для напряжения питания 12 В [3].

В качестве приводов натяжителя и намотчика используются коллекторные электродвигатели постоянного тока RS-380SH-12300 и ZGB37RG соответственно. Рабочий ток электродвигателя RS-380SH-12300 составляет 0,31 А при напряжении питания 12 В. Момент силы на валу составляет 63,9 гр×см. Рабочий ток электродвигателя ZGB37RG составляет 0,6 А при напряжении питания 12 В. Момент силы на валу составляет 1 кг×см.

Для отображения информации используется жидкокристаллический индикатор 16x2, который имеет две строки по шестнадцать символов. Данный индикатор используется с переходником

на I2C интерфейс, что позволяет использовать для управления дисплеем два аналоговых выхода контроллера вместо шести цифровых. Переходник имеет встроенный потенциометр для ручной регулировки контрастности дисплея. Одним из основных преимуществ данного устройства является его способность избавить оператора от необходимости контролировать скорость намотки.

На рисунке 2 представлен алгоритм работы управляющей программы.

СПИСОК ЛИТЕРАТУРЫ

1. Микроконтроллер ATmega328 [Электронный ресурс]. – Режим доступа: <http://www.alldatasheet.com/datasheet-pdf/pdf/392243/ATMEL/ATMEGA328.html>.
2. Шаговый двигатель ПБМГ 200 265 – характеристики [Электронный ресурс]. – Режим доступа: <http://libixur.ru-/2012/11/shagovyj-dvigatel-pbmg-200-265-harakteristiki>.
3. Датчик Холла А3144 [Электронный ресурс]. – Режим доступа: <http://www.alldatasheet.com/datasheet-pdf/pdf/55092-/ALLEGRO/A3144.html>.

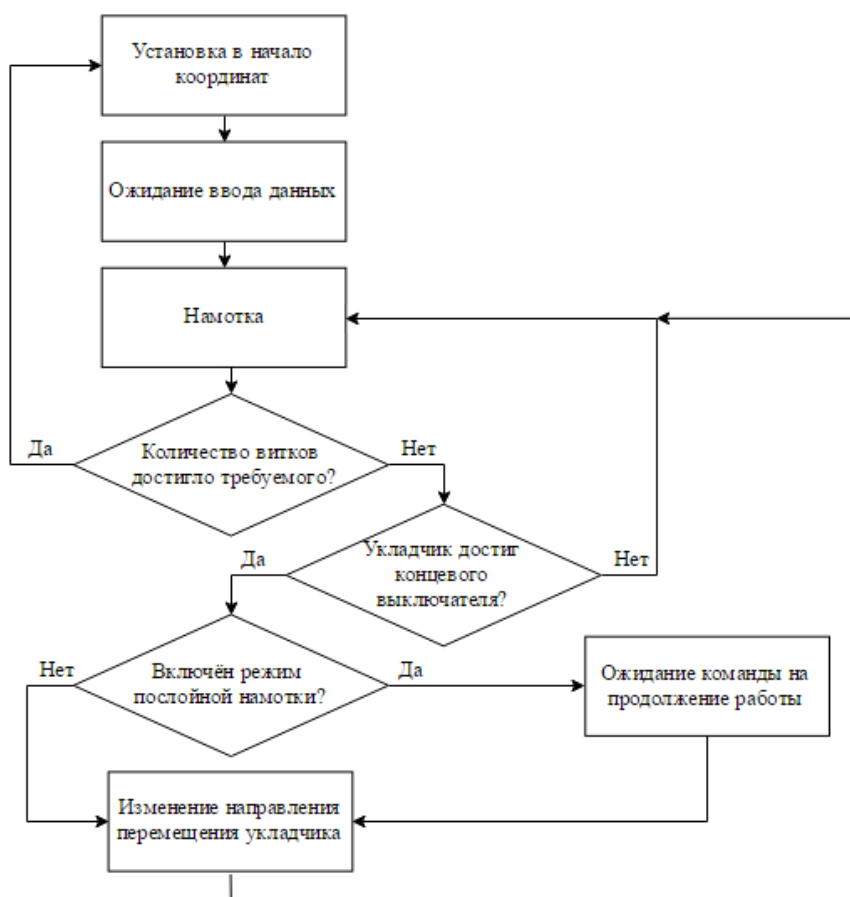


Рис. 2 – Алгоритм работы

ВОПРОСЫ БЫСТРОДЕЙСТВИЯ В СИСТЕМАХ УПРАВЛЕНИЯ ЭЛЕКТРОПРИВОДОМ

Соколов В. И., Хаджинов М. К.

Кафедра электронных вычислительных средств, кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: xrrg.z500@gmail.com, kh_m@tut.by

ПИ- и ПИД-регуляторы широко применяются в сервоприводах различных устройств. Негативное влияние на динамические характеристики систем с данными типами регуляторов оказывает эффект интегрального насыщения. В настоящей работе рассматриваются методы устранения данного эффекта.

ВВЕДЕНИЕ

На данном этапе развития науки и техники широкое распространение получил электропривод. Существует ряд применений, где к качеству переходного процесса предъявляются серьезные, противоречащие друг другу требования. Так от системы ожидается высокое быстродействие, точность регулирования, а также нежелателен колебательный характер переходного процесса. Наилучшим показателем быстродействия обладает регулятор релейного типа, однако установившийся процесс в системе с данным типом регулятора имеет колебательный характер. Для управления электроприводом (двигателем постоянного тока) зачастую применяются пропорционально-интегральный (ПИ), а также пропорционально-интегрально-дифференцирующий (ПИД) регуляторы. Однако, на динамические характеристики систем с данными типами регуляторов оказывает эффект интегрального насыщения.

I. ЭФФЕКТ ИНТЕГРАЛЬНОГО НАСЫЩЕНИЯ

В установившемся режиме работы и при малых возмущениях большинство систем с ПИД регуляторами являются линейными. Однако процесс выхода на режим практически всегда требует учета нелинейности типа "ограничение" которая связана с естественными ограничениями на мощность, скорость, частоту вращения, угол поворота и прочими ограничениями, которые накладываются на исполнительный механизм. Контур регулирования в системе, находящейся в насыщении (когда переменная достигла ограничения), оказывается разомкнутым, поскольку при изменении переменной на входе звена с ограничением его выходная переменная остается без изменений. В это время канал интегрирования продолжает накапливать величину ошибки во времени, которая не оказывает никакого влияния на объект управления. Такая ситуация называется интегральным насыщением (integral windup).[1]

Эффект интегрального насыщения проявляется ухудшением динамических характери-

стик системы управления. В частности, увеличивается перерегулирование, а также время переходного процесса, что проиллюстрировано на рисунке 1 (y - выходная координата, y_{sp} - величина уставки, u - управляющее воздействие, I - интегральная составляющая).

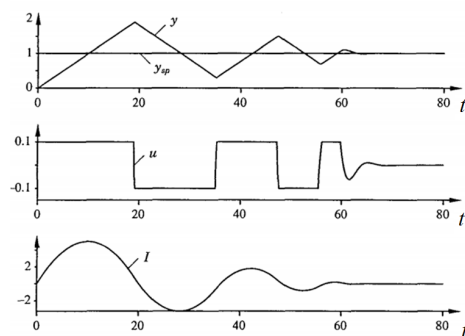


Рис. 1 – Демонстрация интегрального насыщения

II. МЕТОДЫ УСТРАНЕНИЯ ИНТЕГРАЛЬНОГО НАСЫЩЕНИЯ

Основная идея этих методов состоит в том, чтобы не допустить нежелательного роста интегральной составляющей при выходе управления из линейной зоны. Эти способы можно разделить на пять основных групп:

1. условное интегрирование (conditional integration – CI);
2. метод обратного счета (back-calculation – BC);
3. метод задания интегральной составляющей (preloading – PR);
4. ПИД-регулятор переменной структуры (variable structure PID – VSPID);
5. комбинированный метод, сочетающий условное интегрирование с методом обратного счета (CI-BC).

Условное интегрирование. При использовании этого метода интегральная составляющая растет только при выполнении некоторых условий, в противном случае она сохраняет свое значение. Ограничения на рост интегральной составляющей могут быть следующими:

1. интегральная составляющая ограничена заданными величинами;
2. интегрирование останавливается, когда ошибка становится большой, то есть $|e| > e_{max}$, где e_{max} – заданная величина;
3. интегрирование останавливается, когда исполнительный механизм достигает ограничения, то есть $u \neq u_s$;
4. интегрирование останавливается, когда исполнительный механизм достигает ограничения, а знаки ошибки и управления совпадают, то есть $u \neq u_s$ и $e \cdot u > 0$.

К недостатку первых двух способов, которые чаще всего применяются на начальной стадии процесса при выходе на рабочую точку, относится возможное появление статической ошибки. В первом случае, ограничение интегральной составляющей может помешать выходу регулируемой величины на задание. Во втором случае, выход регулятора может «зависнуть» и остаться на таком значении, при котором ошибка больше порогового значения e_{max} . Поэтому для использования этих методов требуется дополнительный анализ объектов управления, что затрудняет их применение. В третьем и четвертом способе эти проблемы решаются, но четвертый способ значительно предпочтительнее третьего, так как в нем интегральная составляющая может изменяться в сторону, противоположную достигнутому ограничению.[2]

Метод обратного счета. Этот способ является альтернативой методу условного интегрирования. Он сводится к пересчету интегральной составляющей регулятора, как только управляющий сигнал выходит за пределы ограничений. Вход интегральной составляющей уменьшается на величину пропорциональную разности между вычисленным значением ПИ- (ПИД-) управления и ограничениями выхода.

Метод задания интегральной составляющей. В этом методе также вводится дополнительная обратная связь, управляющая интегральной составляющей. Эта обратная связь включается при выходе вычисленного значения управления на ограничение. При ее включении выход интеграла динамически стремится к заранее определенной величине i_d . Для управления скоростью убывания интеграла при выходе управления на ограничение применяется параметр α . Недостатком этого метода является необходимость заранее определять настроечные параметры i_d и α . Также нельзя гарантировать, что управление у сойдет с ограничения при стремлении i к i_d .

ПИД-регулятор переменной структуры (VSPID). Этот метод, близкий к методу об-

ратного счета, отличается тем, что пересчет интегральной составляющей, то есть включение обратной связи происходит при выполнении некоторых условий. При выходе вычисленного управления u на ограничения, интегральная составляющая динамически изменяется так, чтобы u оставалось вблизи граничных значений u_{max} или u_{min} . В линейной зоне поведение VSPID алгоритма совпадает с поведением обычного ПИД-регулятора, поэтому использование этого способа не вызывает затруднений на практике. Настойка параметра α также не вызывает затруднения, обычно α выбирается так, чтобы динамика обратной связи в интегральном контуре была в 2...5 раз быстрее, чем динамика замкнутой системы управления. Поскольку управление u в режиме ограничения находится близко к границам линейной зоны, VSPID регулятор достаточно быстро переходит в линейный режим при сходе управления с ограничений.

Комбинированный метод. В этом подходе, аналогично VSPID, к методу обратного счета добавляются некоторые условия из способа условного интегрирования. Вычисления в линии обратной связи для блока интегрирования выполняются, если управляющий сигнал выходит за пределы ограничений, знаки ошибки и управления совпадают. Но в отличие от VSPID во всех режимах ошибка e подается на вход интеграла. Настроечные параметры u и α выбираются аналогично VSPID.

III. ВОЗМОЖНЫЕ АЛЬТЕРНАТИВЫ

Существует класс систем автоматического управления, которые проектируют таким образом, чтобы переходной процесс в них был аperiodическим и как можно быстрее достигал требуемого значения. Такие системы называют системами с минимальным временем переходного процесса или системами, оптимальными по быстрдействию. Такие системы сочетают сильные стороны релейного регулятора (быстродействие), а также преимущества ПИ (ПИД) регуляторов (точность регулирования). Цель дальнейшей работы - исследовать свойства регуляторов данного типа, сравнить их с традиционными схемами.

1. Astrom, K. J. and T. Hagglund. PID Controllers: Theory, Design and Tuning ISA Press. Research Triangle Park, USA 1995.
2. Гуляев С.В., Кузнецов С.И., Малахов В.А., Ольшванг В.Р., Шубладзе А.М. Обзор работы ПИ- и ПИД- регуляторов в режимах интегрального насыщения/Гуляев С.В., Кузнецов С.И., Малахов В.А., Ольшванг В.Р., Шубладзе А.М.//Автоматизация в промышленности. – 2008. №8.– С. 6–9.

СИСТЕМА СБОРА ИНФОРМАЦИИ О СОСТОЯНИИ ЗОНЫ ПОКРЫТИЯ СЕТИ ДЛЯ УПРАВЛЕНИЯ РАЗВИТИЕМ ИНФРАСТРУКТУРЫ ОПЕРАТОРОВ СВЯЗИ

Сорокин А. А.

Кафедра «Связь», Астраханский государственный технический университет

Астрахань, Россия

E-mail: alsorokin.astu@mail.ru

В работе описывается система сбора и обработки информации о состоянии зоны покрытия сети, которую применяют в качестве сигналов обратной связи во время управления эксплуатацией и развитием инфраструктуры оператора связи. В отличие от аналогов предлагаемая система допускает совместное использование в качестве источников информации специализированных датчиков и абонентских устройств. Система предусматривает сбор информации о мобильных устройствах абонентов, что позволяет решать ряд дополнительных задач, в области управления трафиком информационных приложений

ВВЕДЕНИЕ

Управление инфраструктурой оператора связи – сложная задача, требующая учета большого количества факторов. Особую сложность представляет мониторинг и оценка состояния зоны покрытия сети (ЗПС) оператора мобильной связи, поскольку местоположение абонентов может непрерывно изменяться. А между абонентским оборудованием и базовой станцией (БС) сети мобильной связи могут находиться различные препятствия в виде зданий и зеленых насаждений, которые оказывают негативное влияние на условия распространения сигнала. Поэтому для обеспечения равномерности распределения качества предоставляемых услуг на территории ЗПС оператора необходимо выявлять участки, на которых происходит ухудшение уровня мощности сигнала и в зависимости от типа ситуации применять необходимое управляющее воздействие на элементы инфраструктуры сети.

Примером подобных воздействий является изменение углов ориентации антенн, корректировка мощности сигнала, замена отдельных модулей и целиком БС, реализация проектов по созданию новых элементов сетевой инфраструктуры оператора.

Ключевое место в сборе информации о состоянии ЗПС играют системы мониторинга, например [1], которые при помощи датчиков измеряют уровень мощности сигнала, и ряд дополнительных параметров канала между БС и точкой наблюдения. Однако, подобные комплексы имеют ограничения, сужающие возможности использования результатов измерений для управления развитием инфраструктуры сети оператора и решения ряда дополнительных задач в области распределения его ресурсов между клиентами сети.

Цель работы: Описание системы сбора информации о состоянии зоны покрытия сети оператора связи и абонентских терминалов.

I. ОПИСАНИЕ СИСТЕМЫ СБОРА ИНФОРМАЦИИ О СОСТОЯНИИ ЗОНЫ ПОКРЫТИЯ СЕТИ ОПЕРАТОРА СВЯЗИ

Как описывается в работе [2] в настоящее время наиболее широко использование получили системы сбора информации, основанные на следующем принципе. Для непосредственных измерений используется датчик, как правило, в виде специализированного телефона, на котором сохраняется файл результатов измерений в определенных точках ЗПС. После измерений файл передается в локальное хранилище данных о состоянии сети оператора связи, располагающееся, как правило, на персональном компьютере или сервере отдела эксплуатации сети. После обработки файла результаты полученных измерений отображаются на геоинформационной системе в виде карты распределения уровня сигнала. Время и места проведения измерений, как правило, согласуются в отделе эксплуатации и развития сети заблаговременно. Частота проведения измерений составляет от одного до нескольких раз в месяц, в маршруты проведения измерений включаются оживленные автомагистрали и улицы, места скопления людей. В одном региональном отделении оператора часто находятся один – два подобных измерительных комплекса, марки, которых в разных региональных отделениях могут различаться. Для проведения измерений назначаются специалисты и транспортные средства. Ранее в работе [2] отмечалось, что при использовании подобных способов возникают ограничения связанные: с недостаточной частотой проведения измерений (на многих улицах и придомовых территориях городов и участков сельской местности измерения не проводятся); относительно высокой стоимостью программных комплексов; локализованностью баз данных результатов измерений на «территории» региональных отделений; необходимостью использования специальных средств измерений или рекомендованных производителями комплексов телефонных аппаратов. В рамках проводимых исследований

предлагается использовать системы сбора данных о состоянии ЗПС оператора связи дополнить системой, представляющей собой клиент-серверное приложение, состоящее из программы, которая устанавливается на мобильное устройство пользователя, например под управлением Android, IOS или Windows операционной системы и программного обеспечения, устанавливаемого на удаленный Интернет-сервер. Схема организации системы сбора информации представлена на рис. 1.

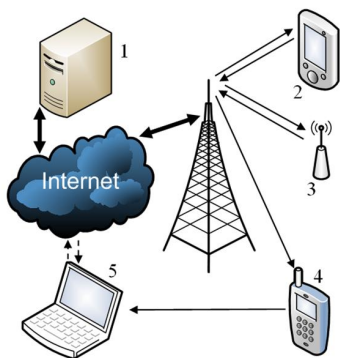


Рис. 1 – Схема организации системы сбора информации о состоянии ЗПС

На рис. 1 приняты следующие обозначения 1 – сервер сбора и обработки данных о состоянии сети, 2 – абонентский терминал пользователя с установленной программой, 3 – стационарный датчик, 4 – терминал с измерительным комплексом другого производителя, 5 – локальный компьютер для сбора данных. Задача программы на абонентском устройстве – взаимодействие с модулями операционной системы, отвечающих за работу с элементами аппаратной части (сетевых интерфейсов и вычислительных систем), после сбора информации она обобщается в виде log-файла и передается через сеть интернет на сервер сбора и обработки данных. Сервер обработки данных агрегирует информацию, получаемую от абонентских устройств и распределяет ее по группам задач, для решения которых она используется.

Описанный принцип реализован в процессе разработки комплекса «Netmetric» [3], используемого для построения карты покрытия сети оператора связи. Данные о состоянии мобильного устройства использованы в процессе разработки программы [4] ориентированной для решения задачи управления потоками трафика во время рассылки пакетов обновлений программного обеспечения. Внешний вид интерфейса программного комплекса «Netmetric», отображающего состояние зоны покрытия приведен на рис. 2., показывающего фрагмент карты распределения уровня мощности сигнала на территории одного из районов г. Астрахань, измерения проводились в период 2015 – 2016 гг. Дополнительно по каждому квадрату возможно получение данных по изменению уровня сигнала за каждый из периодов времени.



Рис. 2 – Фрагмент карты покрытия с распределением уровня мощности сигнала

Применительно к решению задачи управления потоками трафика во время рассылки пакетов обновлений методами теории нечеткого вывода обобщалась информация учитывающая вычислительную способность, энергообеспеченность, качество взаимодействия с сетью оцениваемого мобильного устройства. По результатам оценки формировалась метрика, которая позволяла назначать мобильное устройство на роль мобильного сервера, который самостоятельно проводил рассылку уже полученных пакетов обновлений другим аналогичным устройствам с меньшим значением метрики.

II. ЗАКЛЮЧЕНИЕ

В ходе проведения исследований предложена система сбора информации о состоянии ЗПС оператора связи основанная на дополнении существующих методов мониторинга состояния ЗПС возможностью получения данных с дополнительных статичных датчиков и устройств абонентов. Предложенная система предусматривает агрегирование информации на едином сервере с последующим ее ранжированием для решения задач в области управления сетевой инфраструктурой оператора связи.

1. QualiPoc Android. The premier handheld troubleshooter. Режим доступа: <http://www.swiss-qual.com/en/products/optimization2/qualipoc-android/>
2. Сорокин, А. А. Распределенная измерительная система сети сотовой связи на основе мобильных датчиков / А. А. Сорокин, А. А. Горюнов, Д. С. Марочкин // Датчики и системы. – 2017. – № 3.(212). – С. 16–23.
3. Свидетельство о государственной регистрации программы для ЭВМ № 2016610998 от 20.01.2016 г. МПК (нет) «Система комплексного мониторинга сети мобильной связи NetMetric» / А. А. Сорокин, А. А. Горюнов –2016.
4. Свидетельство о гос. регистрации программы для ЭВМ № 2017612841 от 03.03.2017 г. МПК (нет) «Fuzzy module for calculation the metric of mobile node» / А. А. Сорокин, А. А. Горюнов, А. Э. Ходарова, Э. Р. Тлекова –2017.

ОПТИМИЗАЦИЯ ВРЕМЕНИ ОЖИДАНИЯ ПРИ ДИСПЕТЧЕРИЗАЦИИ ГРУППЫ ЛИФТОВ НА ОСНОВЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Струц Д. А.

Кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: dmitry.a.struts@gmail.com

Рассматривается использование генетического алгоритма для решения проблемы кодирования в диспетчеризации кабин. Предлагаемый метод уменьшает среднее время ожидания и использует более простой подход кодирования, который более дешевый с точки зрения вычислительной стоимости.

ВВЕДЕНИЕ

Элемент управления лифтами является одной из важных проблем вертикальных транспортных систем в зданиях. С точки зрения качества обслуживания два основных параметра для оптимизации, - время ожидания и время в пути.

I. ПАРАМЕТРЫ ВРЕМЕНИ И ТРАНСПОРТНАЯ МОДЕЛЬ

С точки зрения качества обслуживания эффективный алгоритм систем управления лифтовой группой должен посылать кабины лифтов на зарегистрированные вызовы, минимизируя время ожидания пассажиров. Время ожидания - время с регистрации вызова и до достижения кабиной заданного этажа. Время в пути начинается с прибытия лифта на этаж с пассажиром и заканчивается при прибытии на место назначения. Оценка времени поездки является более сложной, поскольку неопределенность из-за зависимости от входящих вызовов и их неизвестных направлений может существенно повлиять на нее. Путешествие - общее время с регистрации вызова до достижения места назначения и рассчитывается как сумма времени поездки (TT) и времени ожидания (WT).

$$JT = WT + TT \quad (1)$$

В зависимости от здания и времени суток можно наблюдать различные потоки движения. Вместо конкретных шаблонов рассматривается общий случайный трафик, пренебрегая другими ограничениями.

II. ГЕНЕТИЧЕСКИЙ АЛГОРИТМ

Генетический алгоритм (ГА) - это алгоритм оптимизации, основанный на естественном процессе эволюции. В ГА популяция состоит из индивидуумов, каждый из которых представляет собой решение для максимизации (минимизации) функции приспособленности. Индивидуумы (или хромосомы) представляют битовые строки. Для процесса эволюции поколений ГА имитирует естественный отбор, мутацию и скрещивание.

При диспетчеризации лифтов, скрещивание применяется как общая замена генов у двух хромосом, гены которых представляют собой номера лифтов. Когда скрещивание применяется к паре, после случайного выбора номера лифта для каждого гена, кабины назначаются на вызовы на этажах. При мутации, для одной хромосомы случайно выбранный ген, который представляет номер лифта, заменяется другим номером лифта. Функция приспособленности для ГА определяется как:

$$f = 1/T_{av} \quad (2)$$

$$T_{av} = \left(\sum_{i=1}^K WT_i \right) / K \quad (3)$$

где K , WT_i и T_{av} представляют количество вызовов, расчетное время ожидания i -го вызова и среднее время ожидания соответственно. Входные параметры системы, которые будут использоваться при вычислении приспособленности, приведены в таблице 1.

Таблица 1 – Параметры системы

Обозначение	Значение
NF	Количество этажей
NC	Количество кабин
PT	Пассивное время
IT	Время межэтажного перекрытия
HC	Этажные вызовы
CF	Этажи расположения кабин

Сумма времени открытия/закрытия двери и времени погрузки пассажиров обозначено как «пассивное время» (PT).

III. СТРУКТУРА ХРОМОСОМЫ

В работе [1], следуя той же схеме, представленной в [2] для кодирования кабин в ГА, Тартан определил хромосому как конкатенированные массивы, где каждый массив представляет состояние присвоения номера кабины вызовам на этажах. Битовая длина хромосом $2M \times [NF - 1]$, где M - количество кабин, делает её громоздкой для вычисления. В [3] хромосома была адаптирована, где каждый ген соответствует этажу, по которым зарегистрирован вызов, отбрасывая

этажи без вызова. Каждое значение бита представляет собой идентификатор кабины, назначенный этому этажу. В этом подходе длина хромосом, но она более практична, чем двоичное кодирование с включением всех этажей. Этот подход уменьшает вычислительные затраты в ГА за счет уменьшения матрицы популяции от $2P \times M \times [NF - 1]$ до $P \times HC$, где P - количество хромосом.

IV. РАСЧЕТ ПРОГНОЗИРУЕМОГО ВРЕМЕНИ ОЖИДАНИЯ

В [4] были продемонстрированы возможные маршруты кабины. Маршруты кабины к этажу, с конечным пунктом назначения выше, чем этаж с вызовом, представлены на рис. 1.

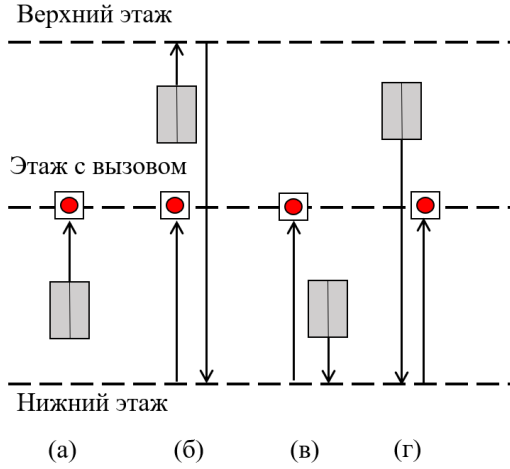


Рис. 1 – Маршруты кабины

Если кабина не движется или поднимается, расчетное время ожидания для i -го этажа с вызовом WT_i рассчитывается в (4), иначе в (5). Здесь n - номер кабины, присвоенный вызову на этаж. Как видно из (4) положение кабины CF_n , относительно позиции этажа с вызовом HC_i и направление движения определяет формулу для расчета. NS_i - минимальное количество остановок между этажом вызова и этажом с кабиной, назначенной на этот вызов. Вызовы регистрируются, и в условиях предварительной нагрузки известны номера кабин. Для выбора решения, ГА

рассматривает этажи с вызовами, которые назначены на эту кабину в NS_i . В обычных системах конечные пункты назначения вызова неизвестны, так как есть кнопки направления вверх и вниз. Значит промежуточный вызов может появиться до того, как кабина достигнет HC_i . Если кабина поднимается вверх, на самый высокий этаж для неё, то следующие вызовы будут назначены на кабину в обратном направлении. Поэтому проверяются первые условия в (4) и (5) и добавляются дополнительные условия. Если в кабина не имеет больше вызовов в своём направлении, то после выхода пассажира, который находился внутри, она может взять пассажира в обратном направлении. Тогда новые ограничения для первых условий в (4) и (5) становятся соответственно (6) и (7).

ЗАКЛЮЧЕНИЕ

Рассматриваемая модель сочетает некоторые свойства предыдущих исследований и улучшает метод, представленный в (Tartan, 2014). Данный метод:

1. Принимает во внимание начальные состояния кабины и условия предварительной нагрузки;
2. Использует более компактную десятичную схему кодирования вместо двоичной кодировки;
3. Улучшает функцию приспособленности для предполагаемого времени ожидания.

1. Tartan, E. O., Erdem, H. and Berkol, A. (2014). Optimization of waiting and journey time in group elevator system using genetic algorithm, Innovations in Intelligent Systems and Applications (INISTA) Proceedings, Italy.
2. Cortes, P., Larraneta, J. and Onieva, L.(2004). Genetic algorithms for controllers in elevator groups: Analysis and simulation during lunchpeak traffic, Appl. Soft. Comput., vol. 4, no. 2, pp. 159–174.
3. Bolat, B., Altun, O., Cortes, P.(2013). A particle swarm optimization algorithm for optimal car-call allocation in elevator group control systems, Appl. Soft Comput., 13 (5) , pp. 2633–2642.
4. Jamaludin, J., Rahim, N. A. and Hew, W.P. (2009). Development of a self-tuning fuzzy logic controller for intelligent control of elevator systems. Eng Appl Artif Intel, 22(8): 1167–1178.

$$T_i = \begin{cases} [HC_i - CF_n] \times IT + NS_i \times PT; & \uparrow HC_i \geq CF_n \\ [(NF - CF_n) + (NF - HC_i)] \times IT + NS_i \times PT; & \downarrow HC_i \\ [(NF - CF_n) + (NF - 1) + (HC_i - 1)] \times IT + NS_i \times PT; & \uparrow HC_i < CF_n \end{cases} \quad (4)$$

$$T_i = \begin{cases} [CF_n - HC_i] \times IT + NS_i \times PT; & \uparrow HC_i \geq CF_n \\ [(CF_n - 1) + (HC_i - 1)] \times IT + NS_i \times PT; & \downarrow HC_i \\ [(CF_n - 1) + (NF - 1) + (NF - HC_i)] \times IT + NS_i \times PT; & \uparrow HC_i < CF_n \end{cases} \quad (5)$$

$$T_i = \begin{cases} [(\downarrow HC_{max} - CF_n) + (HC_{max} - HC_i)] \times IT + NS_i \times PT; & \downarrow HC_{max} = CDF_n(max) \\ \text{if all } \downarrow HC < CD_n; & \end{cases} \quad (6)$$

$$T_i = \begin{cases} [(CF_n - \uparrow HC_{min}) + (HC_i - \uparrow HC_{min})] \times IT + NS_i \times PT; & \downarrow HC_{min} = CDF_n(min) \\ \text{if all } \uparrow HC > CD_n; & \end{cases} \quad (7)$$

СИНТЕЗ УСТРОЙСТВА СОВМЕШНОГО ИЗМЕРЕНИЯ-РАСПОЗНАВАНИЯ С АДАПТАЦИЕЙ К УГЛАМ ПРОСТРАНСТВЕННОЙ ОРИЕНТАЦИИ РАДИОЛОКАЦИОННЫХ ОБЪЕКТОВ

Свинарский М. В., Ярмолик С. Н., Храменков А. С., Зайко Е. В.

Кафедра автоматики, радиолокации и приемо-передающих устройств, Военная академия Республики

Беларусь

Минск, Республика Беларусь

E-mail: mechislav1993@gmail.com, yarmsergei@yandex.ru, xras.tech@mail.ru, zayko.eugene@mail.ru

В работе представлены результаты синтеза устройства совместного измерения-распознавания. Полученное устройство позволяет по измеренным оценкам углов пространственной ориентации цели адаптировать эталонные радиолокационные портреты к условиям наблюдения. Учет пространственной ориентации наблюдаемого объекта повышает эффективность решения задачи их распознавания.

ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Распознавание целей является одной из важных задач, которые должна решать радиолокационная станция в процессе своего функционирования. Под радиолокационным распознаванием принято понимать процесс отнесения исследуемого объекта, задаваемого совокупностью наблюдаемых реализаций принятого сигнала, к одному из взаимоисключающих классов [1]. При этом важно определить тактическое назначение наблюдаемых целей, поскольку такая информация, позволяет судить о намерениях противника, дает возможность оптимального построения атаки и расходования боевого ресурса [1]. По мере развития цифровой техники, а также в связи с расширением круга боевых задач, актуальность вопроса распознавания постоянно возрастает.

Процесс прогнозирования изменяющихся условий наблюдения объектов предполагает учет большого числа взаимосвязанных факторов, когда доступная информация характеризуется разнообразными формами и степенью детерминированности. Применительно к радиолокационным системам в качестве классификационных признаков широко используют радиолокационные портреты объектов наблюдения [2]. Под радиолокационным портретом (РЛП) понимают упорядоченную совокупность комплексных амплитуд отраженного сигнала, распределение квадратов модулей которых по рассматриваемым координатам соответствует распределению мощности отраженного сигнала [2].

Используемые РЛП зависят от распределения отражательных свойств цели вдоль анализируемой координаты пространства наблюдения, а также от ряда информативных и малоинформативных параметров [2]. Одним из параметров, который оказывает существенное влияние на эффективность классификации объектов, является пространственная ориентация объекта наблюдения (Θ) [2, 3]. Пространственную ориентацию радиолокационной цели в системе координат ли-

нии визирования (ЛВ) радиолокатора принято характеризовать совокупностью трех углов пространственной ориентации (УПО): курса, тангажа и крена [3].

Процесс принятия решения о классе наблюдаемого объекта предполагает сопоставление реализации обрабатываемого РЛП с имеющимися эталонными портретами. Следует отметить, что в процессе полета летательного аппарата (ЛА) постоянно изменяются его УПО относительно радиолокационной станции, обуславливая необходимость адаптации эталонных РЛП к текущим условиям наблюдения. Отсутствие информации о текущих значениях УПО ЛА существенно затрудняет решение задачи адаптации.

В процессе исследований была разработана методика оценивания УПО ЛА по имеющейся радиолокационной информации. Однако обеспечение высокой точности оценивания УПО ЛА вынуждает затрачивать значительные временные ресурсы [3]. Необходимость использования компромиссного решения между точностью оценивания УПО ЛА и располагаемым временным ресурсом вынуждает совершенствовать подходы к построению адаптивных систем.

Одним из способов адаптации РЛП к текущим условиям наблюдения, является учет закона распределения оценок УПО цели. В связи с тем, что радиолокатор ведет работу в условиях ограниченного временного ресурса, задачу измерения УПО цели и решения об ее классе необходимо решать совместно.

Целью доклада является синтез структуры устройства совместного измерения-распознавания радиолокационных объектов с адаптацией эталонных РЛП к пространственной ориентации распознаваемой цели.

ОСНОВНАЯ ЧАСТЬ

Одним из способов повышения эффективности современных радиотехнических систем является совместное решение задач радиолокационного наблюдения [4]. Совместное решение

задач измерения параметров и распознавания классов рассмотрено в работах [4, 5]. При синтезе структуры таких радиотехнических систем в условиях априорной неопределенности (в данном случае относительно пространственной ориентации объекта наблюдения) наиболее распространенным является Байесовский подход. Применительно к синтезу радиолокационных систем в условиях априорной неопределенности байесовский подход подразумевает формирование оценок среднего (апостериорного) риска на основе данных наблюдения и минимизацию этих оценок выбором правил решения [4, 5].

Следует отметить, что решение задачи радиолокационного распознавания ведется в условиях параметрической неопределенности относительно пространственной ориентации объекта наблюдения в системе координат ЛВ радиолокатора. Задача отождествления сопровождаемых траекторий с разовыми оценками решена, цель наблюдается и обнаруживается в каждом обзоре. Требуется определить класс сопровождаемой цели.

В качестве критерия синтеза устройства использован минимум апостериорного риска (1), совместного распознавания класса объекта наблюдения и измерения его пространственной ориентации, с простой функцией стоимости принятия решения о классе цели и квадратичной функцией потерь для задачи измерения [4, 5]. В формуле (1): A_l – гипотеза о, объекте l -го класса; $\hat{\Theta}$ – оценка УПО ЛА; ξ – элементы РЛП; $p(\xi)$ – безусловная плотность распределения ξ ; $p(A_g)$ – априорная вероятность появления объекта g -го класса; $\Lambda(\xi|A_g)$ – условное отношение правдоподобия при справедливости гипотезы A_g ; $\Pi_l g$ – стоимости за решение об объекте l -го класса при

условии, что он g -го; Θ – истинное значение УПО ЛА; M – количество классов.

Путем минимизации апостериорного риска (1) получено двухфункциональное решающее правило распознавания класса объекта наблюдения с адаптацией к его пространственной ориентации. Полученное решающее правило включает в себя: оценку пространственной ориентации объекта наблюдения и принятие решение о его классе.

ЗАКЛЮЧЕНИЕ

Синтезировано устройство распознавания объектов наблюдения с адаптацией к углам их пространственной ориентации. Алгоритм работы синтезированного устройства совместно измерения-распознавания определяется двухфункциональным решающим правилом, позволяющим получить оценку УПО ЛА и использовать ее для адаптации РЛП объекта наблюдения с целью распознавания его класса.

1. Фукунага, К. Введение в статистическую теорию распознавания образов / К. Фукунага. – М.: Наука, 1979. –307 с.
2. Курлович, В. И. Основы теории радиосистем: учеб. пособие / В. И. Курлович, С. В. Шаляпин. – Минск : ВА РБ, 1999. –342 с.
3. Ярмолик, С. Н. Повышение точности оценивания ориентации летательного аппарата в интересах адаптации радиолокационных портретов к условиям наблюдения / С. Н. Ярмолик, М. В. Свинарский, А. С. Храменков, Е. В. Зайко. // Доклады БГУИР, 2018. – №5. С. 57–64.
4. Трифонов, А. П. Совместное различие сигналов и оценка их параметров на фоне помех / А. П. Трифонов, Ю. С. Шинаков. – М.: Радио и связь, 1986. – 264 с.
5. Репин, В. Г. Статистический синтез при априорной неопределенности и адаптации информационных систем / В. Г. Репин, Г. П. Тартаковский. – М.: «Советское радио», 1977. –432 с.

$$R(A_l, \hat{\Theta}, \xi) = \frac{1}{p(\xi)} \sum_{g=1}^M p(A_g) \Lambda(\xi|A_g) \left[\Pi_l g + \int_{\Theta} (\hat{\Theta} - \Theta)^T (\hat{\Theta} - \Theta) p(\Theta|\xi, A_g) d\Theta \right], l = \overline{1, M} \quad (1)$$

МОДЕРНИЗАЦИЯ ПРОГРАММНОГО КОМПЛЕКСА ПРОМЫШЛЕННОГО СЕРВОПРИВОДА

Тимахович А. Е., Хаджинов М. К.

Кафедра электронных вычислительных средств, кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: stewart96at@gmail.com, khm@bsuir.by

Рассматриваются возможности внедрения квадратичной обратной связи в контур управления программы контроллера промышленного серийно выпускаемого сервопривода.

ВВЕДЕНИЕ

Система управления современных промышленных сервоприводов, в большинстве своем, имеет классическую трехконтурную структуру подчиненного регулирования с последовательной коррекцией. Каждый вложенный контур «подчиняется» заданию вышестоящего. Начиная с внутреннего, это: контур тока (момента), контур скорости (частоты вращения), контур положения. Соответственно, каждый контур с помощью своего регулятора поддерживает на заданном уровне свою величину.

Как правило, регуляторы данных контуров ограничиваются регуляторами типа ПИ, ПИД, а попыток применения принципов модального регулирования и их внедрения в промышленные серийно выпускаемые сервоприводы пока не предпринималось.

I. ПОСТАНОВКА ЗАДАЧИ

Ставится задача модернизации программы контроллера существующего промышленного сервопривода.

На начальном этапе необходимо разработать желаемый контур управления, и уже на его основе рассматривать возможности его внедрения в программу контроллера серво. Так, конечной целью данной работы решено было получить сервопривод с квадратичной обратной связью по скорости [1].

После анализа данного способа регулирования были выявлены некоторые обязательные требования к промышленным серво, на основе которых хотим получить привод с желаемыми характеристиками.

II. ТРЕБОВАНИЯ

Промышленные сервоприводы имеют внутри датчик тока в контуре тока, а также датчик положения – энкодер. Но для построения контура управления с квадратичной обратной связью по скорости нам необходим сигнал скорости. Сигнал скорости можно получить как от дополнительного датчика, так и с помощью дифференцирования выходного сигнала энкодера.

Промышленные сервоприводы имеют драйвер с контроллером, в который защита програм-

ма алгоритма регулирования и управления серводвигателем. Тогда встает резонный вопрос: а насколько велик объем памяти программ контроллера и существует ли какой-то ее запас? Ведь предполагается, что в программу алгоритма будут внесены корректировки и добавлено что-то новое.

Для организации сервопривода с квадратичной обратной связью понадобится очень малая часть данной памяти, так как модернизация будет производиться на основе основного существующего алгоритма управления лишь с несущественными изменениями.

В любом случае, можно с уверенностью сказать, что контроллер драйвера имеет определенный запас памяти хотя бы на том основании, что иногда возникают ситуации, когда разработчики поставляемого сервопривода обнаруживают баги в работе серво, и тогда программа драйвера модифицируется, а драйверы перепрошиваются. Для таких случаев создатели должны предусмотреть возможность внезапного расширения алгоритма.

III. РЕАЛИЗАЦИЯ

Рассмотрим наиболее благоприятный случай, когда привод имеет встроенный датчик скорости. Структурная схема сервопривода с введением квадратичной обратной связи представлена на рис. 1.

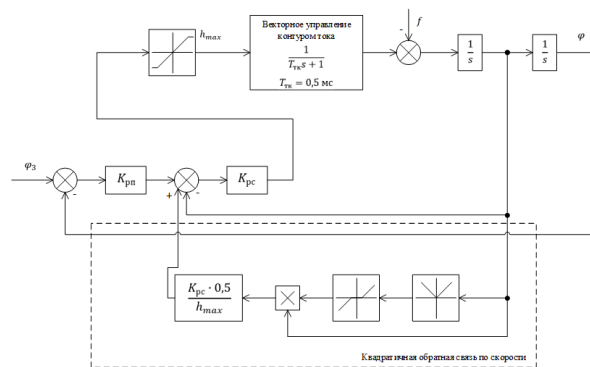


Рис. 1 – Структурная схема контуров управления с квадратичной обратной связью по скорости

Сигнал управления контуром тока подается через ограничитель, соответствующий максимальным силовым возможностям сервопривода.

Объект контура тока представлен в виде апериодического звена с векторным управлением, где $T_{tk} = 0,5$ мс – это время регулирования момента серводвигателя. С помощью сигнала с выхода контура скорости создается квадратичная обратная связь. Коэффициенты усиления контуров положения (K_{rp}) и скорости (K_{rs}) должны параметризоваться в сервоприводе.

Результаты моделирования полученного контура приведены на рис. 2.

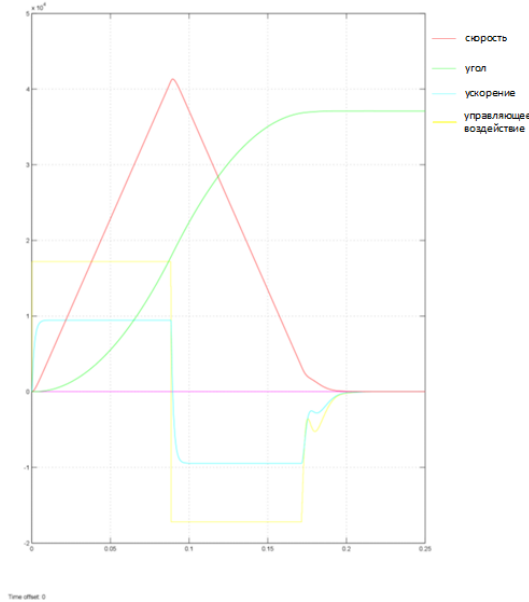


Рис. 2 – Переходные характеристики системы при отсутствии возмущающего воздействия

Как видно из графика, переходный процесс положения плавный и без перерегулирования, угловая скорость изменяется по линейному закону. График ускорения имеет релейный вид с дотягиванием в зоне линейности.

Результаты моделирования системы с учетом возмущающего воздействия представлены на рис. 3.

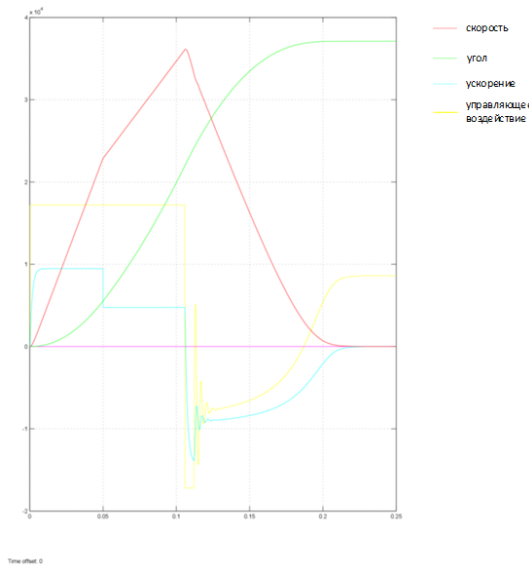


Рис. 3 – Переходные характеристики системы с учетом возмущений

Графики показывают, что переходные процессы остались без перерегулирования, но увеличилось время регулирования.

Моделирование показало эффективность применения квадратичной обратной связи по скорости.

IV. ПРОБЛЕМЫ

Основной трудностью при решении данной задачи является непосредственно внедрение желаемой программы, потому что алгоритмы управления производителей промышленных сервоприводов являются технологиями ноу-хау и получить доступ к исходному коду программы контроллера весьма проблематично. Но сейчас предпринимаются попытки в решении данной проблемы.

ЗАКЛЮЧЕНИЕ

Таким образом, задача получить сервопривод с квадратичной обратной связью на основе существующего промышленного сервопривода является вполне реализуемой. Для ее решения необходимо выполнить следующие условия: получить сигнал скорости от внутреннего скорости, иметь достаточное количество вычислительных средств, которое позволит добавить в исходный алгоритм свои наработки.

Результаты моделирования показали, что при введении в структурную схему управления сервоприводом квадратичной обратной связи по скорости переходные процессы получаются оптимальными по быстродействию, изменяются практически по релейному закону с дотягиванием в линейной зоне.

Однако возникает проблема получения доступа к исходному коду программы контроллера сервопривода.

СПИСОК ЛИТЕРАТУРЫ

1. Хаджинов М.К., Доманов А.Т., Павлова А.В. Сервоприводы с квадратичной обратной связью по скорости. Мн., БГУИР 2017.

СИСТЕМА ПЕРЕМЕЩЕНИЙ УСТАНОВКИ АВТОМАТИЧЕСКОГО КОНТРОЛЯ ОРИГИНАЛОВ ТОПОЛОГИИ

Титко Д. С., Салманзадех Г. Й., Марко А. Ф.

Кафедра высшей математики, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: mmts@bsuir.by

В работе представлена и описана прецизионная система перемещений установки автоматического контроля оригиналов топологии, используемая для достижения при сканировании точного совмещения координатной системы фотошаблона и координатной системы самой установки контроля и с чувствительностью до 10 нм.

Основными операциями контроля качества процессов фотолитографии являются контроль критических размеров элементов топологии, контроль координат элементов топологии и совмещаемости различных слоев интегральной схемы, контроль профиля топологических структур (в основном, для полупроводниковых пластин), а также контроль дефектности топологических структур [1–3].

Если для оценки критических размеров, координат элементов и профиля, как правило, производят выборочный контроль при отработке технологии с участием человека, то при контроле дефектности топологических структур на определенных стадиях технологического процесса производится сплошной контроль, для каждого изготовленного изделия. Особое место в технологической цепочке контроля занимает оборудование автоматического контроля оригиналов топологии на фотошаблонах. Для реализации такой технологии контроля на ГНПО ТМ «Планар» разработаны и изготовлены установки ЭМ 6329 и ЭМ 6729, которые в автоматическом режиме выполняют контроль оригиналов топологии СВИС путем сравнения изображения маски фотошаблона с искусственным изображением, сгенерированным из проектных данных [4].

Принцип работы установки контроля оригиналов топологии основан на сравнении реальной топологии шаблона с его эталонным описанием, полученным из системы автоматизированного проектирования топологии. Установка контроля состоит из оптико-механического устройства, устройства управления и стола оператора с терминалом. Оптико-механическое устройство содержит двухкоординатный стол, позволяющий производить перемещения с чувствительностью 10 нм, механизм ориентации шаблона, оптико-электронный преобразователь с линейным многоэлементным фотоприемником с зарядовой связью, осветитель для контроля шаблонов в проходящем свете и визуального наблюдения в проходящем и отраженном свете, систему автофокусировки, бинокулярный микроскоп для визуального наблюдения, переносной пульт управления.

Устройство управления содержит специализированный анализатор изображений с инженерным пультом, блок управления координатной системой, блок усилителей мощности, одноплатную промышленную ЭВМ канала реального изображения, одноплатную промышленную ЭВМ канала эталонного изображения, рабочую станцию, управляющую циклом установки и осуществляющую связь с оператором, блоки питания, блок автоматики и блок развязки с сетевыми фильтрами.

Для достижения при сканировании точного совмещения координатной системы фотошаблона и координатной системы самой установки контроля используется прецизионная система перемещений, принципиальная структурно-кинематическая схема которой показана на рис. 1.

Высокий уровень совмещения достигается за счет обеспечения возможности получения более полной совместимости координатных систем установок комплекса. Эта совместимость достигается за счет применения однотипных датчиков линейных перемещений, построенных на базе интерферометров с двухчастотным лазером (рис. 1), а также за счет реализации аналогичных алгоритмов управления перемещениями координатного стола. Координатные столики установок контроля построены по единой схеме на базе линейных шаговых двигателей. В цепи обратной связи применяются однотипные интерферометрические датчики линейных перемещений, построенные на базе двухчастотных лазеров. Все это позволяет при автоматическом контроле топологии существенно сократить количество ложных ошибок, возникающих за счет отклонения траекторий перемещений координатных столов генератора изображений и установки автоматического контроля топологии и, тем самым, повысить достоверность контроля. С другой стороны, появляется возможность более точного выхода в зону дефектов при их устранении, что упрощает процедуру ремонта металлизированной маски фотошаблона.

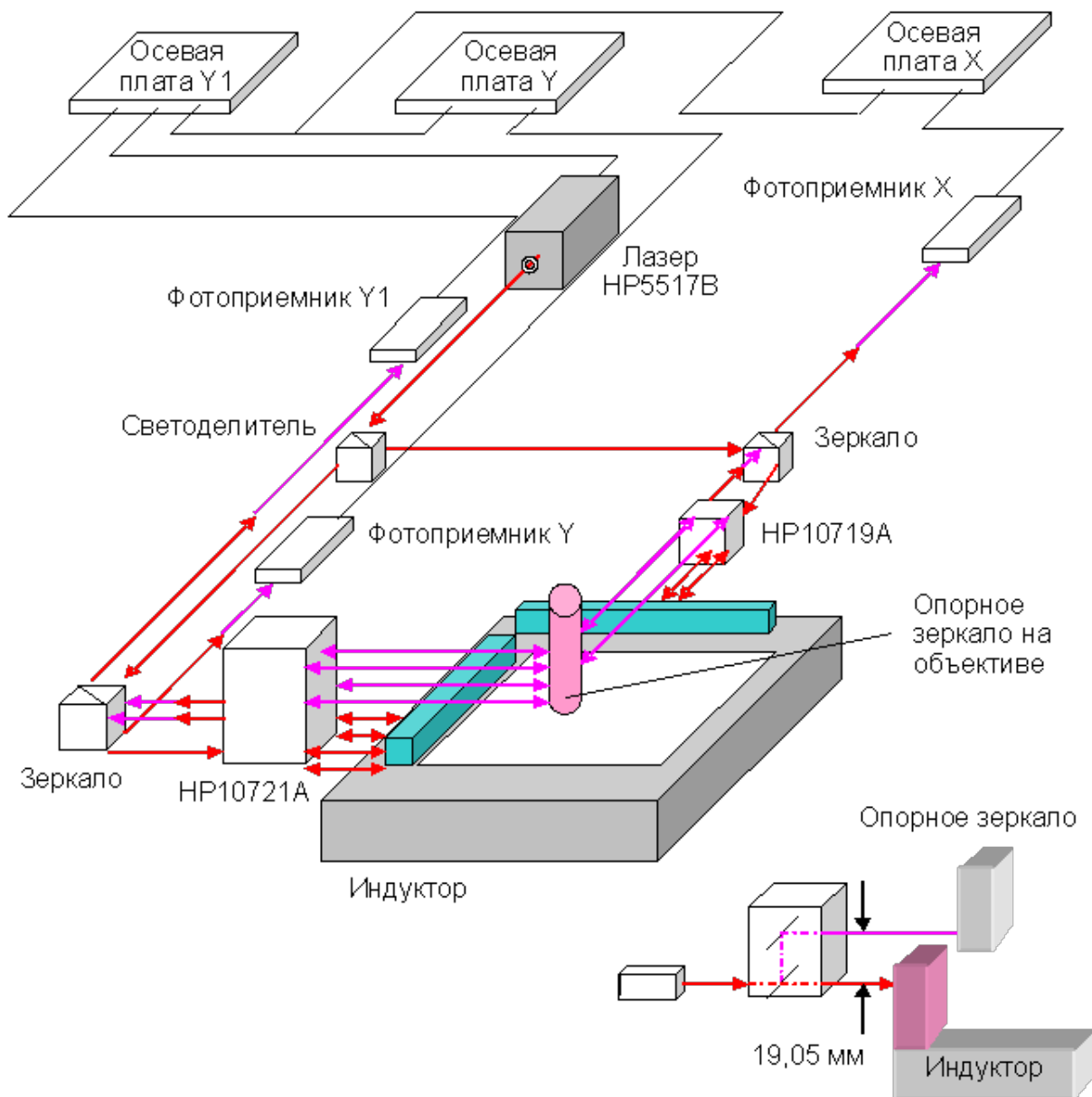


Схема деления луча в интерферометрах на опорный и измерительный

Рис. 1 – Система перемещений установки контроля оригиналов топологии

При этом обеспечивается возможность реализации единых подходов к построению алгоритмов компенсации погрешностей координатных систем. К ним относятся погрешности, связанные с изменениями окружающей среды, с локальными неравномерностями шаблонов, с разбросом параметров при изготовлении составных частей координатных систем.

1. Карпович, С.Е. Системы многокоординатных перемещений и исполнительные механизмы для прецизионного технологического оборудования : моногр. / С.Е. Карпович [и др.] ; под. ред. д-ра техн.

наук, проф. С.Е. Карповича. – Минск : Бестпринт, 2013. – 208с.

2. Аваков, С.М. Автоматический контроль топологии планарных структур : моногр. / С.М. Аваков; науч. ред. С.Е. Карпович. – Минск : ФУАинформ, 2007. – 168 с.
3. Аваков, С.М. Компенсации погрешностей совещения при автоматическом контроле топологии планарных структур СВИС : моногр. / С.М. Аваков // Доклады БГУИР. – 2007. – № 6. – С. 72–77.
4. Оптико-механические комплексы для бездефектного изготовления фотошаблонов 0,35 мкм и 90 нм / С.М. Аваков [и др.] // Фотоника. – 2007. – № 6. – С. 35–39.

АНАЛИЗ И РАЗРАБОТКА МОДУЛЕЙ ДЛЯ ГРУППОВОЙ РОБОТОТЕХНИКИ

Волков А. В.¹, Никифоров А. С.², Чижов И. В.³, Гапанович Ю. А.⁴, Семенович А. С.⁴,
Миранович Р. Б.⁴, Самардак Д. В.⁴

¹Кафедра системного анализа и компьютерного моделирования, ²кафедра биофизики, ³кафедра физической электроники и нанотехнологий, ⁴кафедра информатики и компьютерных систем, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: andrei@cybergizer.com, sky92033@live.com, f_kknopsikk@mail.ru, ura345710@gmail.com, alexser1999@gmail.com, h, roman.miranovich@gmail.com, samardak.dasha31@gmail.com

В докладе представлен аналитический обзор существующих групповых минироботов. В результате выявлены и указаны достоинства и недостатки исследованных подходов. Предложена новая конструкция систем группового миниробота отвечающая требованиям модульности, вычислительной мощности, богатства аппаратной периферии.

ВВЕДЕНИЕ

Групповая, или роевая робототехника — это направление научных исследований в области достижения коллаборативного взаимодействия между автономными роботами в процессе решения поставленной задачи, выполнение которой не достижимо каждым из роботов в отдельности [1]. Исследования в данной области могут выполняться как с помощью имитационного моделирования коллективного поведения [2], так и с применением физических роботов [1]. Последний способ позволяет учитывать экспериментальные условия, которые могут быть упущены в симуляциях, что формирует особый интерес в научном сообществе к данному подходу. Особое распространение в академических исследованиях получили роботы относительно малых размеров (менее 7.5 на 7.5 см², именуемые далее как минироботы) в связи с малыми размерами, относительной простотой реализации и дешевизной производства, что способствует проведению экспериментов для больших групп роботов (например, 1024 робота в случае Kilobot [3]). В настоящее время предложено множество платформ групповой робототехники [1]. Направление новых научных разработок сводится к преодолению недостатков разработанных решений, однако сопряжено с множеством трудностей рассматриваемых далее.

I. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

Локальная коммуникация рассматривается исследователями как основополагающая особенность робототехнических систем роевого взаимодействия [4]. Каждый агент взаимодействует только с соседними агентами и хранит лишь локально обретенную информацию. Большинство существующих минироботов используют инфракрасные (ИК) передатчики и приемники в силу того, что данный способ позволяет осуществлять локальный обмен сообщениями, а также идентифицировать близ расположенных роботов [1].

Недостатком ИК канала передачи данных является подверженность помехам, шумам и ошибкам, а также низкая скорость передачи данных.

Максимально возможные бортовые вычислительные возможности являются необходимым условием для запуска разного рода алгоритмов коллективного поведения в режиме реального времени. Однако большинство существующих минироботов (например, [3, 5, 6]) построены на базе 8-битных микроконтроллеров, ограниченных в скорости обработки данных, количестве встроенных периферийных устройств и требующих дополнительных вычислительных циклов для простейших операций таких как умножение в сравнении с 32-битными микроконтроллерами.

Множество существующих минироботов характеризуются отсутствием модульности, что затрудняет последующее обновление аппаратного и программного обеспечения, а также препятствует проведению исследований в широком спектре экспериментальных условий. Использование платы выполняющей роль бортового компьютера и связующей все остальные модули миниробота посредством последовательных интерфейсов (таких как I²C) способствует преодолению данного недостатка.

Некоторые групповые минироботы, такие как TinyTeRP [7] и Kilobot [3] способны измерять лишь расстояние до близрасположенного миниробота и не способны определять азимутальное направление к нему, что ограничивает спектр сценариев коллективного поведения и значительно усложняет сами алгоритмы группового взаимодействия.

Целью настоящей работы является разработка модулей для групповых минироботов, призванных устранить выявленные недостатки. Разработанный подход должен поддерживать модульный принцип построения аппаратного обеспечения робота, удовлетворять требованиям вычислительной эффективности, обеспечивать воз-

возможность эффективной беспроводной коммуникации с соседними минироботами.

II. ПРЕДЛАГАЕМЫЙ ПОДХОД

Модуль обработки данных и коммуникации

Модуль обработки данных и коммуникации (МОДК) является центральным звеном в предлагаемой архитектуре миниробота, обеспечивая подключение систем робота по интерфейсам I²C, SPI. Тип микроконтроллера (тактовая частота, архитектура, встроенная периферия такая как аналогово-цифровой преобразователь, цифро-аналоговый преобразователь и т. д.) определяют типы пригодных к применению датчиков на минироботе. Принимая во внимание данные аспекты, в качестве МОДК выбрана система на кристалле ESP32 [8]. Альтернативой для ESP32, выступает система на кристалле ESP8266 [8], которая может быть использована в задачах требующих меньшей вычислительной мощности, а также меньшего встроенной периферии. Для реализации беспроводного обмена сообщениями выбрана библиотека построения mesh-сети painlessMesh [9] с поддержкой микросхем ESP32/ESP8266.

Модуль обнаружения соседа

Для реализации принципа локальной коммуникации, каждый миниробот должен знать идентификаторы соседних минироботов, что позволяет их адресацию в mesh-сети. Для определения соседних роботов разработан ИК модуль представленный на рис. 1.

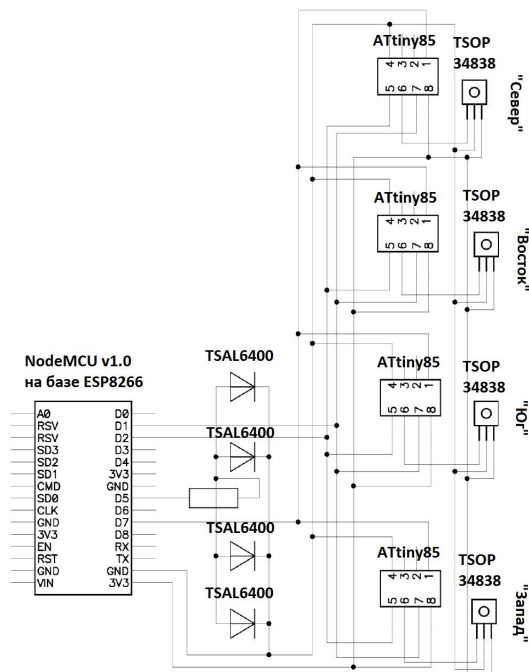


Рис. 1 – Принципиальная схема модуля обнаружения соседа

Передача идентификационного номера осуществляется по NEC протоколу (частота несущей 38 кГц). При этом используется ИК диод TSAL6400 подключенный к цифровому выводу. Для приема модулированного ИК сигнала ис-

пользуются фотоприемники TSOP34838 преобразующих принятый сигнал в цифровой. Микроконтроллер ATtiny85 осуществляет декодирование цифрового сигнала с сохранением значения в буфер. Приемники и передатчики размещены в направлениях "Север" "Юг" "Запад" "Восток" для обеспечения покрытия четырех сторон света (рис. 1). Использование четырех микроконтроллеров ATtiny85 позволяет вести прием ИК сигнала во всех направлениях одновременно. МОДК подключен к плате обработки данных и коммуникации по I²C интерфейсу. Вынесение обработки принимаемого ИК сигнала в отдельную систему позволяет снять данного рода вычислительную нагрузку с МОДК и обеспечить модульность.

III. ЗАКЛЮЧЕНИЕ

Предложен модуль обработки данных и коммуникации на базе системы на кристалле ESP32. Возможность локальной коммуникации обеспечивается посредством передачи данных по mesh-сети. Направления будущих исследований включают разработку исполнительных устройств (таких как модуль передвижения) и модулей сбора сенсорной информации, а также совершенствование разработанного модуля обнаружения ближайших соседей.

Работа поддержана фондом «Science Around Us Foundation» (Лондон).

IV. СПИСОК ЛИТЕРАТУРЫ

1. Hamann, H.: Swarm Robotics: A Formal Approach. Springer (2018)
2. Pinciroli, C., Trianni, V., O'Grady, R., Pini, G., Brutschy, A., Brambilla, M., et al. (2012). ARGoS: A modular, parallel, multi-engine simulator for multi-robot systems. Swarm Intelligence, 6(4), 271–295. ISSN 1935-3812.
3. Rubenstein, Michael, et al. "Kilobot: A low cost robot with scalable operations designed for collective behaviors." Robotics and Autonomous Systems 62.7 (2014): 966-975.
4. Brambilla, M., Ferrante, E., Birattari, M., & Dorigo, M. (2013). Swarm robotics: A review from the swarm engineering perspective. Swarm Intelligence, 7(1), 1–41. ISSN 1935-3812.
5. S. Yim and S. Kim, "Origami-Inspired Printable Tele-Micromanipulation System," IEEE Int. Conf. on Robot. and Autom., Seattle, WA, 2015, pp. 2704–2709.
6. F. Arvin, J. Murray, C. Zhang, and S. Yue, "Colias: An autonomous micro robot for swarm robotic applications," Int. J. Adv. Robot. Syst., vol. 11, no. 1, 2014.
7. A. P. Sabelhaus, D. Mirsky, L. M. Hill, N. C. Martins, and S. Bergbreiter, "TinyTeRP: A Tiny Terrestrial Robotic Platform with Modular Sensing," IEEE Int. Conf. on Robot. and Autom., Karlsruhe, Germany, 2013, pp. 2600–2605.
8. A. Maier, A. Sharp and Y. Vagapov, "Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things," 2017 Internet Technologies and Applications (ITA), Wrexham, 2017, pp. 143-148.
9. PainlessMesh [Electronic resource] Mode of access: <https://gitlab.com/painlessMesh/painlessMesh>. Date of access: 25.09.2018.

ПРЕИМУЩЕСТВА И ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ НЕМИНИМАЛЬНО-ФАЗОВОГО КОЭФФИЦИЕНТА ОТРАЖЕНИЯ В ЗАДАЧАХ ШИРОКОПОЛОСНОГО СОГЛАСОВАНИЯ СОПРОТИВЛЕНИЙ

Янцевич М. А., Дубовик И. А.

Кафедра автоматики, радиолокации и приёмо-передающих устройств, Военная академия Республика
Беларусь

Минск, Республика Беларусь

E-mail: yantsevich.mikhail@mail.ru

Показан аналитический подход к разрешению ограничений на согласование сопротивлений, с помощью рационального использования параметров коэффициента отражения.

В настоящее время во всем мире наблюдается быстрое развитие радиоэлектронных систем: систем сотовой и радиорелейной связи, радионавигации и радиолокации, телевидения, быстродействующих систем передачи данных, радиоизмерительных комплексов и т.д. При проектировании радиоэлектронных устройств, входящих в состав этих систем (таких как усилители, преобразователи и умножители частоты, активные фильтры, антенные устройства и др.), важное значение имеет решение задач широкополосного согласования, а также коррекции формы частотной характеристики радиоэлектронных устройств. Суть этой задачи в общем случае состоит в согласовании комплексных импедансов генератора и нагрузки для максимизации передаваемой мощности в заданном диапазоне частот. Для решения указанной задачи используются реактивные согласующие цепи (СЦ). Классические методы широкополосного согласования, основанные на аналитической теории Фано-Юлы, предполагают решение задачи синтеза согласующих цепей в два этапа – сначала аппроксимация частотных характеристик синтезируемой цепи, затем – реализация. Метод изложенный в [1][2], обладает потенциалом для реализации СЦ применительно к сложным комплексным нагрузкам. Данный подход разрешает ограничения на согласования нагрузки путём изменения параметров аппроксимирующей функции. Используемая функция Баттерворта в [1] имеет единственный варьируемый параметр – коэффициент уровня передачи (рисунок 1).

Главной характеристикой СЦ, определяющей передачу мощности, является коэффициент передачи по мощности $K(\omega)$. Кроме него, цепь характеризуется также коэффициентом отражения. Существенное влияние на результат синтеза СЦ оказывает представление коэффициента отражения минимально и неминимально-фазовой функцией. Последняя обладает значительным

преимуществом и имеет вид [1]:

$$\rho(s) = \pm \frac{(\delta^n + \sum_{i=1}^n (-1)^i b_i s^i)}{a_0 + \sum_{i=1}^n a_i s^i} \quad (1)$$

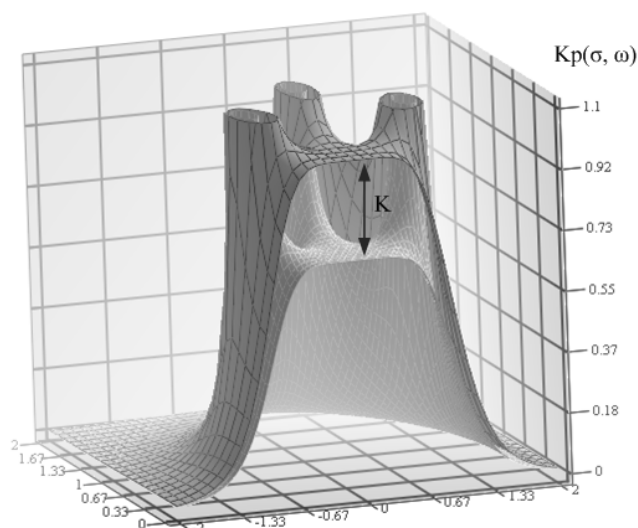


Рис. 1 – Сечение поверхности функции передачи Баттерворта

Коэффициент уровня передачи имеет вид:

$$K = 1 - \delta^{2n} \quad (2)$$

где $\delta \in [0,1]$ в соответствии с [1];

Однако указанный параметр можно изменять в пределах $(-1;1)$, что в значительной мере повышает возможности согласования комплексных нагрузок. При этом выполняется ещё одно важное условие - сохранение вида частотной характеристики. Таким образом изменения диапазона варьируемого параметра коэффициента отражения раскрывает потенциал аппроксимирующей функции в задачах согласования.

Для подтверждения теоретических доводов, необходимо решить практическую задачу. Анализ возможностей согласования с использованием неминимально-фазового коэффициента отражения с изменением параметра δ в диапазоне $(-1;1)$, продемонстрируем на низкочастотной RCL нагрузке, схема которой представлена на рисунке 2. В качестве функции передачи выберем функцию Баттерворта 3-го порядка. Нормированные параметры нагрузки соответственно равны: $R_H=1$; $C_H=1$; $L_H=5$

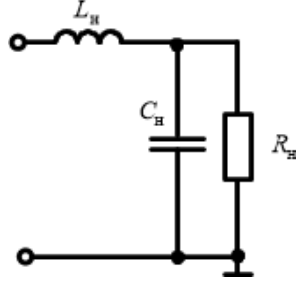


Рис. 2 – Схема низкочастотной RCL - нагрузки

В общем, задачи синтеза СЦ в соответствии с [1] сводится к решению системы уравнений (3). Последняя, включает в себя уравнения: определяющие вид частотной характеристики, ограничения на пределы согласования, а также условие принадлежности функции к классу ограниченно-вещественных функций.

$$\begin{cases} a_0 = 1; \\ 2a_2 - a_1^2 = 0; \\ a_2^2 - 2a_1a_3 = 0; \\ a_3 = b_3 = 1; \\ 2\delta^3 b_2 - b_1^2 = 0; \\ b_2^2 - 2b_1b_3 = 0; \\ a_3 + b_3 - CR(a_2 - b_2) = 0; \\ \frac{R(a_2 - b_2)}{a_1 + b_1 - CR(a_0 - \delta)} - L \geq 0; \\ a_1a_2 - a_3 \geq 0 \end{cases} \quad (3)$$

Далее используя систему z-параметров [1] получаем функцию выходного сопротивления:

$$Z_{\text{вых}}(s) = \frac{1.601}{0.399 + 0.399s}; \quad (4)$$

Ниже представлены результаты синтеза схемы СЦ, а также вид частотной характеристики СЦ с нагрузкой.

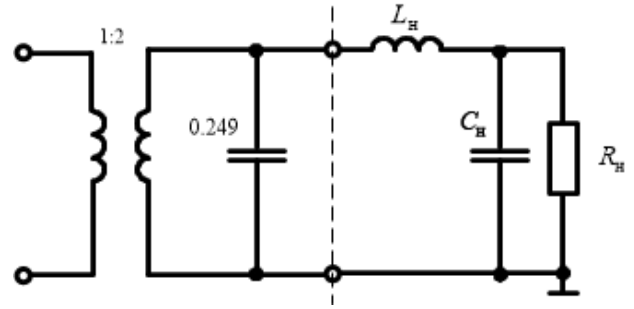


Рис. 3 – Схема ШСУ с нагрузкой

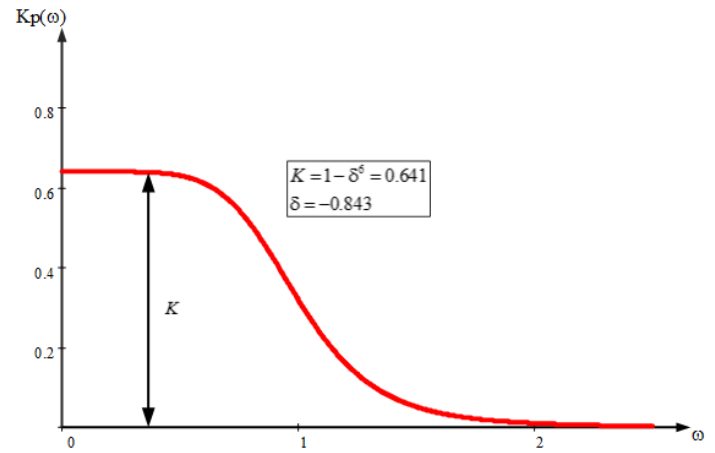


Рис. 4 – Частотная характеристика передачи мощности

Диапазон значений $\delta \in (0,1)$ накладывает более жёсткие ограничения на параметры нагрузки. Для рассматриваемой нагрузки с использованием неминимально-фазового коэффициента отражения где $\delta \in [0,1)$, максимальное значения параметра $L_H=2$. Параметр $\delta \in [-1,0)$ позволяет значительно расширить диапазон значений принимаемых параметрами нагрузки. В приведённом примере значения параметра L можно увеличить в 2.5 раза. Полученный результат имеет значение как теоретическое так и практическое значение в синтезе СЦ.

ПРОБЛЕМЫ СИСТЕМ УПРАВЛЕНИЯ ТУРБОГЕНЕРАТОРНОЙ УСТАНОВКОЙ

Ярохович А. А., Хаджинов М. К.

Кафедра электронных вычислительных средств, Кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: aleksey.yarokhovich.kuvo@gmail.com

В данной работе были рассмотрены существующие системы управления турбогенераторной установкой, выявлены существующие проблемы и поставлены задачи для поведения дальнейшей работы.

ВВЕДЕНИЕ

Качество генерируемой энергии в основном зависит от системы автоматического управления паротурбинных генераторов, которая осуществляет первичное регулирование частоты сети с помощью автоматического регулятора частоты вращения (АРЧВ) турбин. Ввод новых турбогенераторных установок с современными системами автоматического управления в течение длительного периода крайне незначителен. Поэтому решение данной проблемы является трудно реализуемым. Целью исследования является анализ современных методов управления паровыми турбогенераторными установками, методов расчёта элементов системы управления и защиты турбины в целом.

I. АВТОМАТИЧЕСКИЕ СИСТЕМЫ УПРАВЛЕНИЯ СОВРЕМЕННЫХ ПАРОВЫХ ТУРБОГЕНЕРАТОРНЫХ УСТАНОВОК (ТГУ)

Автоматические системы управления современных паровых турбогенераторных установок (ТГУ) — это сложные комплексы различных устройств регулирования и защиты установки. Система управления турбогенераторной установкой обеспечивает четкую и безотказную

работу, автоматически поддерживает заданные режимы работы турбоагрегата, осуществляет измерение и отображение технологических параметров. Кроме того, системы управления обеспечивают необходимые безопасность и надежность всего агрегата при аварийных ситуациях.

Структура системы управления турбогенераторной установкой зависит от типа, мощности и принципа работы турбины. Конструктивное исполнение узлов турбогенераторов и устройств этих систем у разных производителей также различно. Но в любой системе управления используются принципиально одинаковые способы решения поставленных задач. Поэтому можно рассматривать некоторые общие принципы построения САУ паровых турбин, не разбирая подробно их конструктивного выполнения [1-2].

Общая для большинства системы автоматического управления скоростью вращения турбины структурная схема представлена на рис.1. Здесь на вход в систему в сумматоре складываются значение установки угловой скорости вращения турбины s от задатчика и передаточная функция возмущающего воздействия от систем реакторного отделения W_{po} , вызывающие незначительные колебания давления в главном паровом коллекторе. Определим общую передаточную функцию САУ турбины [3].

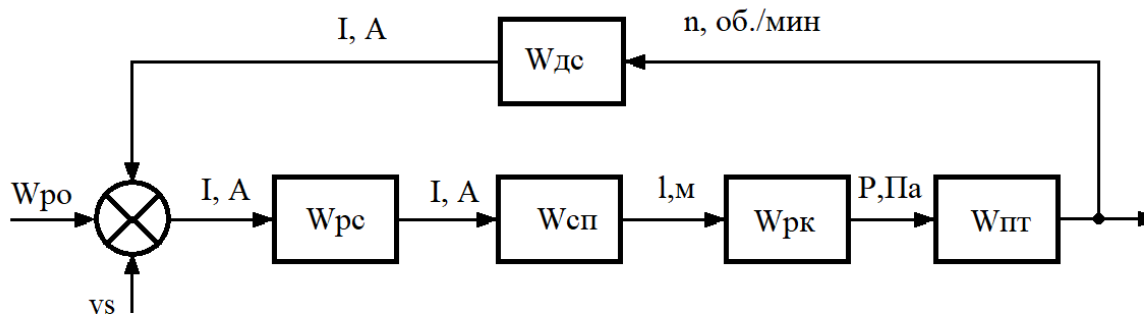


Рис. 1 — Структурная схема линейной модели САУ турбины: РС — регулятор скорости вращения ротора турбины; СП — электрогидравлический следящий привод; РК — регулирующий клапан; ПТ — паровая турбина; ДС — датчик угловой скорости

Общая передаточная функция САУ турбины:

$$W_T(p) = \frac{W_{PC} \cdot W_{СП} \cdot W_{ПК} \cdot W_{ПТ}}{1 + (W_{PC} \cdot W_{СП} \cdot W_{ПК} \cdot W_{ПТ} \cdot W_{\Delta C})} \quad (1)$$

$$W_T(p) = \frac{77.91 \cdot p^5 + 855.8 \cdot p^4 + 3334 \cdot p^3 + 5932 \cdot p^2 + 4875 \cdot p + 1500}{1 + (0.01 \cdot p^8 + 0.201 \cdot p^7 + 1.45 \cdot p^6 + 5.52 \cdot p^5 + 12.22 \cdot p^4 + 16.26 \cdot p^3 + 18.82 \cdot p^2 + 5.52 \cdot p + 1)} \quad (2)$$

II. ПРОБЛЕМЫ И ЗАДАЧИ

Существующие проблемы систем управления турбогенераторной установкой:

1. Поддержание постоянной частоты вырабатываемой электроэнергии. Основной задачей системы управления паровой турбины является поддержание постоянной скорости вращения ротора турбины независимо от изменяющейся нагрузки. Поэтому первичная задача системы управления заключается в удержании частоты в допустимых пределах при нарушении баланса активной мощности. Однако, частота до номинального значения не восстанавливается, что обусловлено статизмом систем управления.
2. Поддержание баланса мощности. Повсеместное потребление электроэнергии задаёт случайные отклонения потребления, а следовательно, и неплановое нарушение баланса мощности. Изменение нагрузки (выходной мощности) во время работы паровой турбины может существенно повлиять на ее производительность и эффективность[4].
3. Долгое время запуска генераторной установки. Для запуска или изменения выходной мощности электростанций с ТГУ может потребоваться много часов, если не дней.
4. В режиме холостого хода генератора переходный процесс носит явно выраженный колебательный характер, особенно при больших значениях коэффициента усиления САУ.

Задачи исследования:

1. С целью ускорения процессов регулирования и избаления от колебаний частоты в режиме холостого хода заменить ПИ-регуляторы контуров управления давлением и мощности на модальные регуляторы.
2. Для устранения влияния чистого запаздывания на качество процессов регулирования

перестроить структуру системы управления в виде прогнозирующего наблюдателя. Запаздывание из контура управления устраняется и остаётся лишь в контуре оценивания.

3. Использовать в наблюдателе упрощенную модель с встроенным компенсатором перекрёстных связей и с запаздываниями, перенесёнными на выход.
4. Ввести в систему управления ограничения на управляющие переменные и попытаться подавить режим автоколебаний квадратичными обратными связями по производной регулируемой величины. Сигналы производных формировать в модели наблюдателя.
5. Провести сравнительный анализ моделей существующих САУ ТГУ с моделью полученной в результате исследования.

III. ВЫВОДЫ

Существующие АСУ ТГУ частично решают выявленные проблемы, но для лучшего их решения следует: изменить применяемые ПИ-регуляторы на модальные с целью приближения модели СУ к реальным процессам, ввести ограничения на уровне управления и квадратичной обратной связью по производной добиться регулирования на пределах ограничений.

1. Новосёлов В. Б. Разработка методов исследования и совершенствования электрогидравлической системы регулирования и защиты паровых теплофикационных турбин и их элементов. Екатеринбург, 2014.
2. Bosch Rexroth «Гидравлика высокого давления в современных газо- и паротурбинных системах» (RE 09 722/07.90).
3. Ефремова Т. А., Мартюшев Д. Н. Расчёт характеристик системы автоматического управления мощностью энергоблока АЭС // Молодой ученый. — 2015. — №22.5. — С. 39-42. — URL <https://moluch.ru/archive/102/23651/> (дата обращения: 22.09.2018).
4. Ротов П.В. Повышение энергетической эффективности работы ТЭЦ и городских систем теплоснабжения. Москва, 2011. № 9. С. 16–20.

О РАЗРАБОТКЕ СИСТЕМЫ РЕКОМЕНДАЦИЙ И ПРОКЛАДКИ МАРШРУТОВ К ТУРИСТИЧЕСКИМ ОБЪЕКТАМ

Заблоцкий В. В., Рудикова Л. В.

Кафедра современных технологий программирования, Учреждение образования «Гродненский государственный университет имени Янки Купалы»

Гродно, Республика Беларусь

E-mail: {rudikowa, viktorzablotsky}@gmail.com

Излагаются общие подходы к реализации веб-приложения с микросервисной архитектурой на платформе JVM, которое предназначено для создания рекомендаций и поиска маршрутов к туристическим местам. Приводится основная функциональность программного обеспечения, связанного с разработкой указанного веб-приложения.

ВВЕДЕНИЕ

Многие люди часто совершают поездки и путешествия с различными целями. Любое путешествие невозможно без планирования, т.к. необходимо решить, в какое место ехать, где остановиться, найти жилье, а также решить, как проводить досуг. Относительно недавно люди, собираясь в путешествие, тратили много времени и усилий на поиск необходимой информации, нередко собирая туристическую литературу и обзванивая знакомых или туристические агентства. Однако с развитием Интернета стало возможно планировать путешествие не выходя из дома. В современном мире путешественник имеет всю необходимую под рукой благодаря Интернет-ресурсам, например, сервис по бронированию отелей, туристический гид или сервис для поиска попутчиков. Благодаря прорыву в разработке современных веб- и мобильных интерфейсов на выбор доступны десятки тематических приложений, в той или иной степени заслуживающих внимания пользователя. Туристическая сфера действительно огромна, в ней нашли свою нишу не только гиганты IT индустрии, такие как Google и Apple, но и множество небольших компаний и самостоятельных разработчиков.

Таким образом, в настоящее время возможно спланировать каждый аспект своего путешествия, используя различные веб-сайты и мобильные приложения. Для сомневающихся путешественников существуют не только справочники со статическим набором тех ли иных объектов, но и системы, специально предназначенные для подбора рекомендаций, исходя из предпочтений пользователя, а также системы, помогающие прокладывать эффективные маршруты между этими объектами. Такие системы на сегодняшний день являются крайне востребованными как среди обычных пользователей, так и разработчиков.

I. ЭТАПЫ РАЗРАБОТКИ ПРИЛОЖЕНИЯ

Работу над реализацией приложения можно разбить на следующие этапы: проектирование инфраструктуры приложения, учитывая тот факт, что приложение состоит из нескольких микросервисов; проектирование каждого микросервиса; проектирование базы данных для каждого микросервиса; проектирование клиентского сервиса; реализация микросервисов и логики взаимодействия между ними; реализация баз данных и логики взаимодействия микросервисов с ними; реализация логики взаимодействия клиентской части приложения с серверной; реализация логики взаимодействия приложения со сторонними API; развертка приложения в виде кластера в облачных сервисах.

Некоторые аспекты разработки приложения включают в себя имплементацию авторизации через протокол OAuth 2, реализацию взаимодействия микросервисов через шину сообщений, работу с картой мира в интерфейсе приложения, а также интеграционное и юнит-тестирование приложения.

II. ОБЩИЕ ПОДХОДЫ К РЕАЛИЗАЦИИ ПРИЛОЖЕНИЯ

Рассмотрим основные компоненты архитектуры веб-приложения для создания рекомендаций и прокладки маршрутов к туристическим объектам.

Данное приложение разработано по принципам микросервисной архитектуры, что означает наличие нескольких слабо связанных между собой сервисов, выполняющих отдельные задачи и образующих единую систему, которая взаимодействует с интерфейсом приложения по принципам REST. Крайне важно правильно разграничить предметную область каждого сервиса для того, чтобы минимизировать их связанность. Это необходимо, чтобы уменьшить время отклика на запросы к серверу, ведь если данные сервисов будут тесно связанными, нужно будет реализовывать механизм транзакций с помощью протокола двухфазного коммита или отдельного

REST API. Если же данные слабо связаны и механизм транзакций не нужен, общей практикой для реализации взаимодействия сервисов является обмен сообщениями через Message Queue с достаточной гарантией целостности данных, например с помощью RabbitMQ.

Предлагаемое приложение состоит из трех (четырёх, если считать клиентскую часть) основных сервисов: API-сервис, Geo-сервис и сервис аутентификации. Кроме основных, в виде отдельных модулей представлены API Gateway - единая точка доступа к микросервисам и балансировщик запросов, а также конфигурационный сервис, к которому подключаются все остальные сервисы при запуске.

Приложение может быть развернуто в изолированной среде как кластер и использовать собственную подсеть. Для этого используются такие средства, как Docker и Kubernetes. Сервисы, обернутые Docker-контейнерами могут быть размещены как локально в виртуальной машине с помощью Minikube, являющимся частью Kubernetes, так и на физических платформах и контролироваться средствами Kubernetes. При этом как Docker, так и Kubernetes предлагают средства для организации сети и мониторинга состояния развернутого кластера. Более того, существуют облачные сервисы, позволяющие размещать кластеры, такие как Amazon AWS.

Серверная часть приложения реализована на платформе JVM с использованием языка программирования Kotlin и фреймворков Spring Boot, Spring Cloud и отдельных модулей Spring Framework. Клиентская часть разработана с помощью фреймворка Vue.js на языке Typescript, использует карты OpenStreetMap и размещена в среде Node.js.

Рассмотрим кратко каждый из сервисов приложения и его клиентскую часть. Точкой входа, через которую проходят все запросы к приложению, является API Gateway. Этот модуль представляет собой прокси-сервер, перенаправляющий запросы к нужному сервису, исходя из контекста запроса. Кроме того, этот модуль имеет также механизм балансировки запросов, что крайне полезно при горизонтальном масштабировании приложения.

Конфигурационный сервис служит для задачи раздачи параметров сервисов в виде yaml-файлов. Все сервисы приложения должны подключиться и получить свои параметры у сервиса конфигурации, а значит он должен запускаться первым в кластере. Данный сервис позволяет хранить конфигурационные файлы в одном месте, при этом имеет механизм разделения файлов на профили.

Auth-сервис предназначен для аутентификации пользователей в приложении и реализует данный функционал с помощью средств Spring Security и протокола OAuth 2. В общем случае, после успешной аутентификации сервис возвращает клиенту JWT-токен, в котором содержится служебная информация, дающая доступ к определенному ресурсу приложения. Все, что нужно для доступа к защищенным ресурсам - отправить в заголовке запроса этот токен. Следует отметить, что данный сервис выполняет лишь аутентификацию, т.е. проверяет истинность введенных пользователем имени и пароля. Авторизация, или проверка прав пользователя, выполняется каждым микросервисом независимо друг от друга.

API-сервис - главная часть приложения, где реализована основная бизнес-логика. Здесь происходит обработка запросов, не связанных с географическими вычислениями, такими как редактирование настроек пользователей, подбор рекомендаций и иные функции. Этот сервис реализован в классическом MVC-стиле и следует принципам REST архитектуры. В качестве базы данных была выбрана нереляционная СУБД MongoDB, в связи с ее простотой и одновременно мощностью, а также удобным способом хранения данных в формате JSON.

Микросервисом разрабатываемого приложения является также геосервис, в котором реализуется обработка запросов на вычисление локаций по заданным координатам, вычисление маршрутов и импорт данных из публичной базы OpenStreetMap. Также, как и API-сервис, данный сервис использует REST для коммуникаций с клиентом и следует паттерну MVC, уровень данных в котором представляет поисковая база Elasticsearch, выбранная благодаря встроенному механизму обработки геозапросов.

III. Выводы

Таким образом, разработанное приложение предназначено для широкого использования людьми, планирующими путешествие. Веб-система предоставляет удобный интерфейс с картой, позволяющей наглядно получить информацию о туристических объектах в выбранном месте и составить удобный маршрут и график посещения этих мест.

В дальнейшем в приложение могут быть интегрированы новые сервисы, предназначенные для поиска отелей и каршеринга, а также разработана мобильная версия под ОС Android.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Заблоцкий, В. В. О разработке системы рекомендаций и прокладки маршрутов к туристическим объектам / В. В. Заблоцкий //

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ РЕЕСТРА БЛОКОВ ТРАНЗАКЦИЙ ДЛЯ ПОСТРОЕНИЯ РАСПРЕДЕЛЕННЫХ СИСТЕМ

Жилинская Е.Р., Кардаш И.П., Захарьев В.А.

Кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {katyazhilinskaya, ivan.kardash11}@gmail.com, zahariev@bsuir.by

Данная статья посвящена исследованию технологии реестра блоков транзакций для построения распределенных систем (блокчейн). В ней представлены различные типы систем, подробно рассмотрена архитектура, основные консенсусные алгоритмы, а также возможности внедрения технологии как в коммерческих, так и в некоммерческих целях.

ВВЕДЕНИЕ

Реестр блоков транзакций (блокчейн) — выстроенная по определённым правилам непрерывная последовательная цепочка блоков, содержащих информацию. Чаще всего копии цепочек блоков хранятся на множестве разных компьютеров независимо друг от друга. Концепция технологии блокчейн похожа на базу данных, за исключением того, что взаимодействия с ними различаются. Блокчейн — это, по существу, распределенная база данных или публичная книга всех транзакций или цифровых событий, которые были выполнены и распределены между участвующими сторонами в одноранговых сетях.

Блокчейн имеет две основные особенности:

1. Блокчейн является общедоступной. Любой может просмотреть ее в любое время, поскольку она находится в сети, а не в одном учреждении, которому поручено поддерживать и вести запись.

2. Блокчейн также шифруется, она использует шифрование с использованием открытых и закрытых ключей, чтобы гарантировать её безопасность.

Существует семь принципов проектирования для создания программного обеспечения, услуг, бизнес-моделей, рынков и организаций на основе блокчейн: сетевая целостность, распределение нагрузки, ценность как стимул, безопасность, приватность, защищенность прав и вовлеченность.

В основном существует три типа блокчейнов: частная блокчейн, консорциум блокчейн и публичная блокчейн. Частный или консорциум блокчейн связаны с ограниченной средой, такой как компания, группа компаний или одна определенная цепочка создания стоимости, в то время как публичная блокчейн поддерживает тип блокировки без разрешения.

I. АРХИТЕКТУРА БЛОКЧЕЙН

Элементы архитектуры могут варьироваться в зависимости от того, какие типы блокчейна используются. Базовыми элементами являются:

Блок. Блокчейн облегчает высокораспределенную регистрацию для записи транзакций, перенос их на определенный узел в сети и организацию их во времени.

Данные постоянно записываются в сеть через файлы, называемые блоками.

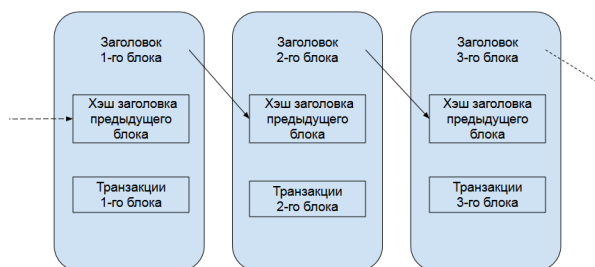


Рис. 1 – Упрощенный пример цепи блоков

Каждый блок состоит из заголовка блока и блока. Заголовок блока состоит из трех наборов метаданных блока.

Цифровая подпись. Для создания транзакции блокчейн требуется цифровая подпись для аутентификации транзакции.

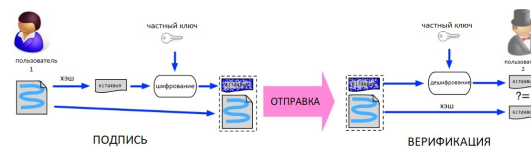


Рис. 2 – Цифровая подпись, используемая в блокчейн

Типичная цифровая подпись включает в себя две фазы: фазу подписи и фазу проверки.

Распределенная сеть. Взаимодействия между пользователем по цепочке блоков в основном используют распределенную сеть, в которой каждый пользователь представляет узел, на котором установлен клин-блок-клин.

Консенсус в сети. Проведение транзакции требует принятия и проверки всеми пользователями в сети, обычно называемой консенсусом. Но каждый узел имеет другой вид состояния всей сети.

Для решения этой проблемы необходим распределенный механизм.

Как правило, существует три основных консенсусных алгоритма, которые могут быть применены:

– Доказательство работы (PoW)

Алгоритм консенсуса является наиболее широко используемым алгоритмом в блокчейн-цепочке. Он был введен биткойном и предполагает, что все коллеги голосуют с их «вычислительной властью», решая случаи и строя соответствующие блоки.

– Доказательство доли (PoS)

Алгоритм направлен на замену существующего способа достижения консенсуса в распределенной системе; вместо решения этот узел, который генерирует блок, должен обеспечить доказательство того, что он имеет доступ к определенному количеству монет до того, как он будет принят сетью.

– Делегированное доказательство доли (DPoS)

Основное различие между PoS и DPoS заключается в том, что PoS является прямым демократическим процессом, в то время как DPoS является представительной демократией - заинтересованные стороны избирают делегатов для создания и проверки блока.

II. ПРИМЕНЕНИЕ ТЕХНОЛОГИИ БЛОКЧЕЙН

Первый блокчейн, биткойн, был разработан для совершенствования системы финансовых приложений. А в настоящее время широко развивается и для некоммерческих услуг, таких как системы голосования для правительственных дел. В связи с этим сферы, использующие блокчейн, можно разделить на два сектора: коммерческие и некоммерческие приложения.

1. Коммерческие приложения

Цифровая платежная система. Это в основном основная функция биткойна как цифровой валюты. Рождение биткойна вызвало эволюцию и нарушение традиционных платежных систем, управляемых банками или другими финансовыми организациями. Схема цифровой валюты включает в себя как новую децентрализованную платежную систему, так и новую.

Смарт-договор. В принципе, смарт-контракт - это компьютерное приложение, которое может автоматически выполнять коммерческие транзакции и соглашения. Он также обеспечивает выполнение обязательств всех сторон в контракте без дополнительных расходов посредника.

Страхование. Любой ценный актив или имущество, которое трудно реплицировать или

уничтожить, можно зарегистрировать в блокчейне. Она может проверять право собственности и отслеживать историю транзакций. Crowdfunding. В настоящее время увеличение числа стартапов реализует криптографические маркеры и протоколы блокчейн в качестве средства поиска своих предприятий. Идея заключается в том, чтобы использовать платформы обратного преобразования, основанные на технологии blockchain, устраняя необходимость в посреднической третьей стороне.

2. Некоммерческие приложения

Распределенное управление. Наиболее распространенное использование блокчейна в органах управления осуществляется в виде нотариуса. Применение блокчейна к нотариальному удостоверению обеспечивает конфиденциальность документа, а также тех, кто для кого производится сертификация. Еще одной формой службы управления, которая также была принята блокчейн, является система онлайн-голосования или электронного голосования.

Распределенное хранилище. Эта концепция была реализована в индустрии здравоохранения и музыки. Для связанных со здоровьем приложений блокчейн предоставляет структуру для хранения медицинских данных или электронных медицинских записей, чтобы их можно было анализировать, но оставаться конфиденциальными.

С другой стороны, в музыкальной индустрии блокчейн применялся для поддержки всеобъемлющей и точной распределенной базы данных прав собственности на музыку.

Распределенный интернет вещей. Использование интернет вещей также представляет некоторые большие проблемы. Один из них связан с централизованной экосистемой, известной также как парадигма клиент-сервер. Хотя эта модель десятилетиями подключала общие вычислительные устройства и будет продолжать поддерживать малые сети интернет вещей, как мы их видим сегодня, она не сможет реагировать на растущие потребности огромных экосистем интернет вещей будущего.

III. СПИСОК ЛИТЕРАТУРЫ

1. Технология Блокчейн и децентрализованное управление: является ли государство все еще необходимо? / Атзори, Марселла // -2015.
2. Кросби, Майкл Технология Блокчейн: Beyond bitcoin. Прикладные инновации 2. - 2016, - С. 6-10.
3. Кастро, Мигель, Лисков, Барбара. Практическая византийская отказоустойчивость. -1999. - С. 173-186.
4. Женг, Зыбин. Проблемы и возможности блокчейн-цепи: обзор. -2016.

Секция «Интеллектуальные
информационные технологии»

Section «Intellectual information
technologies»

ИНТЕЛЛЕКТУАЛЬНАЯ ПРОГРАММА ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ ЭКСПЛУАТАЦИОННОЙ РАБОТЫ НА ПАССАЖИРСКОЙ ЖЕЛЕЗНОДОРОЖНОЙ СТАНЦИИ

Аксёничков А. А.

Кафедра «Управление эксплуатационной работой и охрана труда», Белорусский государственный университет транспорта
Гомель, Республика Беларусь
E-mail: aksen-72@tutl.by

В статье рассматривается применение интеллектуальной программы имитационного моделирования поездных ситуаций для принятия управленческих решений дежурным по железнодорожной пассажирской станции. Приведен расчет эффективности применения данной интеллектуальной программы.

ВВЕДЕНИЕ

Принятие управленческих решений, требующих оценки влияния большой величины транспортного потока с изменяющимися характеристиками, а также изменяющиеся характеристики устройств обслуживания можно учесть с помощью метода имитационного моделирования. Метод имитационного моделирования не имеет ограничений на сложность описываемых объектов и любой фактор может быть учтен, позволяет исследовать сложные системы (которыми являются железнодорожные станции) с учетом связей между их инфраструктурными элементами и внешней средой [1, 2]. При использовании метода имитационного моделирования имеется возможность подготовки универсальных программ для конкретных технологических процессов. Эффективность работы пассажирской железнодорожной станции зависит в первую очередь от того, насколько правильно поставлено управление процессами по приему, отправлению пассажирских поездов и маневровой работе на станции. Выполнение графика движения поездов зависит от действий дежурного по железнодорожной станции (ДСП). Принимаемые ДСП решения зависят в первую очередь от профессиональной подготовки, опыта работы, функционального состояния. ДСП для накопления опыта требуется несколько лет, процесс этот протекает в значительной мере интуитивно. В памяти ДСП накапливаются возникающие ситуации и рациональные решения к ним. Этот подход можно отнести к «методу проб и ошибок». Как показывает практика, специфика работы ДСП включает в себя перечень действий в нестандартных ситуациях, направленных на выполнение графика движения поездов при безусловном обеспечении безопасности движения. В связи с тем, что нестандартные и аварийные ситуации влекут за собой сильный стресс, эмоциональное перенапряжение ДСП, все это влияет на правильность принятия решения.

I. ВОЗМОЖНОСТИ ИНТЕЛЛЕКТУАЛЬНОЙ ПРОГРАММЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Для уменьшения влияния человеческого фактора на принимаемые решения в нестандартных ситуациях, повышения безопасности движения поездов и маневровой работы на станции возникает острая потребность в создании интеллектуальной системы по моделированию поездной ситуации. Основной целью имитационного моделирования поездной ситуации является обеспечение и повышение эффективности и безопасности функционирования пассажирской железнодорожной станции. При пропуске одного пассажирского поезда, следующего со сменной локомотива, ДСП должен принять не менее пяти оперативно-управляющих решений – выбор пути приема поезда, определение времени: открытие входного сигнала, открытия маневрового сигнала для выезда поездного локомотива из под состава в депо, открытия маневрового сигнала, для пропуска поездного локомотива из депо под состав, открытия выходного сигнала по отправлению. А если к поезду производится прицепка или отцепка пассажирских вагонов, то количество операций увеличивается в несколько раз. Все это ведет к психофизиологической нагрузке ДСП и соответственно к принятию им оптимально-правильных решений. В нестандартных и аварийных ситуациях необходимо принимать решения в стрессовой ситуации, ограниченности времени, эмоциональном напряжении, это сказывается на правильности принятия решения. В современном мире для уменьшения психофизиологической нагрузки на ДСП все больше внедряются автоматизированные системы управления, интеллектуальные программные комплексы, информационные технологии и др. Целью интеллектуальной программы имитационного моделирования поездных ситуаций для ДСП, является обеспечение принятия правильных решений в нестандартных и аварийных си-

туациях при приеме и отправлении поездов, выполнении маневровой работы, а именно:

- моделирование маршрутов приема/отправления поездов с учетом времени занятия стрелочных горловин и железнодорожных участков приближения (с учетом интервалов прибытия-отправления, враждебности основных и вариантных маршрутов);
- моделирование маршрутов маневровых передвижений для выполнения размена локомотивов, прицепки/отцепки вагонов, подформирования составов поездов (с учетом интервалов занятия стрелочных горловин);
- увязка поездной и маневровой работы с действующим графиком движения поездов (с учетом занятости путей железнодорожной станции);
- предложение различных вариантов поездной работы с аналитической оценкой и предложение оптимального (с учетом поездной ситуации на подходах к железнодорожной станции и информации от ДНЦ).

Внедрение интеллектуальной программы имитационного моделирования поездных ситуации позволит:

- уменьшит влияние человеческого фактора на принимаемые решения;
- повысить безопасность движения поездов при поездной и маневровой работе;
- повысить достоверность принятия решений по оптимизации использования инфраструктуры пассажирской станции (приемо-отправочных путей), что повлияет на уменьшение непроизводительных эксплуатационных расходов;
- производить оптимизацию поездной и маневровой работы на станции в пассажирском движении, что повлияет на снижение топливно-энергетических ресурсов;
- производить построение прогнозного суточного плана-графика железнодорожной станции в пассажирском движении с учётом различной эксплуатационной нагрузки на станцию, технологии обслуживания пассажирских поездов и маневровой работы.

II. ЭФФЕКТИВНОСТЬ ИСПОЛЬЗОВАНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ ПРОГРАММЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

Эффективность (соотношение между достигнутым результатом и использованными ресурсами (ISO 9000:2015)) [3] интеллектуальной программы имитационного моделирования поездных ситуации можно определить по сокращению затрат времени, связанных с эксплуатацией маневровых локомотивов при отцепке, прицепке вагонов к пассажирскому составу

$$\Delta E = E_{ek} - E_z, \quad (1)$$

где E_{ek} – годовая экономия эксплуатационных затрат от внедрения интеллектуальной программы имитационного моделирования поездных ситуации, руб; E_z – затраты, связанные с внедрением данного программного продукта, руб.

$$E_{ek} = 365 \cdot M_{man} \cdot \Delta t_{man} \cdot n_{op} \cdot e_{l-ch}, \quad (2)$$

где M_{man} – среднесуточное количество маневровых локомотивов, используемых на пассажирской железнодорожной станции по прицепке, отцепке вагонов; Δt_{man} – сокращение непроизводительного времени при выполнении маневровой работы, ч; n_{op} – среднесуточное количество прицепок, отцепок к пассажирским поездам; e_{l-ch} – стоимость маневрового локомотиво-часа, руб.

Рассмотрим на примере станции Минск-Пассажирский эффективность интеллектуальной программы имитационного моделирования поездных ситуации: затраты на разработку программного продукта предположительно составят 65 000 рублей; среднесуточное количество маневровых локомотивов используемых на станции Минск-Пассажирский по прицепке, отцепке вагонов ($M_{man} = 1$); сокращение непроизводительного времени при выполнении маневровой работы ($\Delta t_{man} = 0,03$), ч; среднесуточное количество прицепок, отцепок к пассажирским поездам ($n_{op} = 106$); стоимость маневрового локомотиво-часа ($e_{l-ch} = 69,96^*$), руб [4, 5].

* Приложение 3, таблица 3.1 [4]. Локомотиво (тепловозо)-час маневровой работы – 57,16. Коэффициент пересчета за 1 квартал 2018 года – 1,224.

$$E_{ek} = 365 \cdot 1 \cdot 0,03 \cdot 106 \cdot 69,96 = 81202,$$

$$\Delta E = 81202 - 65000 = 16202.$$

Применение интеллектуальной программы имитационного моделирования поездных ситуации дает положительную эффективность в первый год эксплуатации. Использование данного продукта возможно не только ДСП, но и инженерно-техническими работниками станции в разработке оперативных и перспективных планов работы пассажирской железнодорожной станции.

1. Лещинский, Е. Имитационное моделирование на железнодорожном транспорте. – Пер. с польск / Е. Лещинский. – М. : Транспорт, 1977. – 176 с.
2. Максимей, И. В. Имитационное моделирование на ЭВМ / И. В. Максимей. – М. : Радио и связь, 1988. – 232 с.
3. ISO 9000:2015 Системы менеджмента качества. Основные положения и словарь / Международный стандарт. – М. : «Русский Регистр», 2015. – 53 с.
4. Методические рекомендации по расчету экономических параметров, позволяющих оценить эксплуатационные расходы по технологическим операциям услуг железнодорожного транспорта общего пользования / приказ от 20.12.2016 № 376Н.

ОПРЕДЕЛЕНИЕ ЗЛОКАЧЕСТВЕННЫХ ОПУХОЛЕЙ ГРУДНОЙ ЖЕЛЕЗЫ НА ГИСТОЛОГИЧЕСКИХ СНИМКАХ

Бабеня В. И.

Кафедра веб-технологий и компьютерного моделирования, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: shreiderwalter@gmail.com

Классификация гистологических снимков на данный момент является актуальной задачей. В данной статье описан метод использования сверточных нейронных сетей, пригодный для решения практических задач, связанных с определением злокачественных опухолей на гистологических снимках. Также описано несколько алгоритмов классического компьютерного зрения, имеющие различные качественные характеристики и производится их сравнение.

ВВЕДЕНИЕ

В соответствии с заключением Международного Агентства Раковых Исследований, Всемирной Организации Здравоохранения, 8.2 миллиона смертей были вызваны раком в 2012 году и согласно их прогнозам до 2030 года ожидается ещё 27 миллионов смертей. В частности, рак груди - один из наиболее распространенных типов рака среди женщин. Смертность от рака груди одна из самых высоких в сравнении с другими типами раковых заболеваний. Гистопатологический анализ является крайне времязатратной узкоспециализированной процедурой, успешность которой напрямую зависит от опыта специалиста и подвержена влиянию таких факторов как человеческая усталость и недостаток внимания. Поэтому остро встает вопрос о необходимости создания методов компьютерной диагностики, с целью снизить нагрузку на специалистов патологов, отсеивая наиболее очевидные доброкачественные участки снимков, и таким образом освобождать экспертов для куда более трудоемких для диагностики случаев. Таким образом, вполне естественно, что многие усилия были сосредоточены на анализе гистопатологических снимков грудного рака, в частности автоматической классификации таких снимков на доброкачественные и злокачественные.

I. ОПИСАНИЕ ПРИМЕНЯЕМЫХ МЕТОДОВ

Изображения, приведенные в базе данных, сгенерированы из срезов образцов биопсий и окрашены с помощью гематоксилина и эозина. Они подготовлены в соответствии со стандартной процедурой парафинирования, используемой активно в международной медицинской практике. Основная идея заключается в возможности сохранения структуры тканей, что позволяет наблюдать за ней в простой микроскоп. Вся процедура подготовки образцов состоит из следующих шагов: фиксация образца, его дегидратация, очищение, встраивание и последующая зачистка. После подготовки образцы по-

крываются стеклом. После чего анатомопатолог определяет области с опухолями при помощи микроскопа. Финальный результат диагностики выставляется опытным патологистологом и подтверждается проверочным тестом, таким как иммуногистохимический анализ.

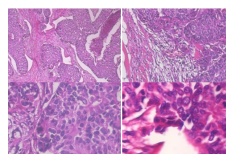


Рис. 1 – Одно изображение злокачественной опухоли (дуальная карцинома) при различных увеличениях: (a) 40x, (b) 100x, (c) 200x, (d) 400X

Гистологические снимки тканей могут быть охарактеризованы при помощи двух подходов. Первый основывается на неявной сегментации изображения, с целью получить структурные свойства рисунка, такие как форма ядер, степень прорастания тканей и так далее, второй подход основывается на глобальной текстурной презентации рисунка. Эти представления конечного изображения были использованы мной как конечные классификаторы для обучения. Они включают в себе также классификаторы часто описываемые в литературе, такие как локальные бинарные шаблоны (LBP), CLBP, локальное квантование фазы (LPQ).

II. НЕКОТОРЫЕ МЕТРИКИ ПРИМЕНЯЕМЫЕ В ПОДХОДЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Оператор LBP состоит из вычисления и распределения бинарных шаблонов в близлежащей окружности каждого пикселя. Близлежащая окружность охарактеризована радиусом R и количество соседей P . Основная идея - пересечь соседние пиксели вместе с центральным: каждому из соседей присваивается значение 1 если яркость текущего пикселя больше или равна яркости центрального пикселя, в противном случае присваивается значение 0. Таким образом, для каждого пикселя бинарный шаблон получается из его окружения. В итоге всего 2^P различных

бинарных шаблонов можно получить. Значение LBP в пикселе p получается путем вычисления скалярного произведения между бинарным кодом и вектором степеней двойки:

$$LBP(p) = \sum_{i=0}^{P-1} 2^i \cdot \delta(f(q_i) - f(p)) \quad (1)$$

где $f(q_i)$ и $f(p)$ это серые уровни пикселей q_i и p соответственно, и δ это функция Кронекера. Гистограмма LBP может быть использована как текстурный декриптор. Стоит иметь в виду, что некоторые шаблоны которые идентичны вплоть до одной или нескольких ротаций могут не иметь одинаковые LBP коды: например, 10000000 и 01000000 имеют соответственно 255 и 128 значения LBP. Этого можно избежать при ротации инвариантного LBP следующего вида: каждый шаблон ротируется P раз и сохраняется наименьшее значение LBP. При помощи этой модификации 10000000 и 01000000 имеют одинаковые LBP значения.

III. КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Использованная мной, для решения описанной задачи, нейронная сеть содержит в себе три сверточных слоя, каждый из которых оборачивается в слои активации и субдескриптации. Первый сверточный слой содержит 96 фильтров, размером $3 \times 7 \times 7$. Слой далее оборачивается в выпрямляющий слой линейного оператора (ReLU), основанный на пулинговой функции максимума, он принимает на вход максимум 3×3 области, содержащие двухпиксельные полоски. Второй сверточный слой содержит 256 фильтров, размером 5×5 . Оборачивается он в аналогичные слои тем, что описаны выше. Итоговый, третий сверточный слой представляет собой множество из 384 фильтров, каждый размером 3×3 . В довершение описанной модели, добавляются два связующих слоя, каждый из которых содержит по 512 нейронов, каждый смежуется с ReLU слоем и исключаяющим слоем, с коэффициентом исключения равным 0.5. Каждый выводящий слой считает дискретное вероятностное распределение всех результатов по всем слоям. Оценить потери при обучении каждого такого слоя можно при помощи функции потерь:

$$L(x, y) = \sum_i (y_i \log p_i(x)) \quad (2)$$

где $p_i(x)$ это вероятность того, что входящий x будет помечен как l_i и y это истинное распределение, такое что $\sum_i y_i = 1$.

Определим финальную погрешность в нашей задаче как взвешенную сумму погрешностей доброкачественности / злокачественности и степени приближения:

$$C = w_{bm}L_{bm} + w_{mag}L_{mag} \quad (3)$$

где w_{bm} и w_{mag} это погрешности для злокачественности и приближения соответственно. Как я упоминал выше, я использовал $w_{bm} = w_{mag} = 0.5$, другие значения могут повлиять на конечные результаты, но проверить это можно только эмпирически.

IV. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

Как показала практика, мешая различные изображения с разных уровней увеличения не ухудшает точность конечного предсказания, а наоборот немного увеличивает его. Само собой это объясняется расширением множества за счет новых степеней увеличения и аффинных преобразований. Для изначально поставленной задачи удалось достичь точности классификации в среднем в 86.7%, что лучше представленного выше подхода более чем на 1.5%. С другой стороны в таблице видно, что распределение точности предсказания по разным степеням увеличения изображений - существенное, что объясняется неравномерным распределением данных в тестовой и обучающей выборках.

V. СПИСОК ЛИТЕРАТУРЫ

1. George, Y. M. Remote computer aided breast cancer detection and diagnosis system based on cytological images / Y. M. George//IEEE -2014. -№ 3.
2. Zhang, Y. Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles / Y. Zhang//Vision Appl. -2013. -№ 7.
3. Zhang, Y. One-class kernel subspace ensemble for medical image classification / Y. Zhang//EURASIP. -2014. -№ 17.
4. Ciresan, D. Mitosis detection in breast cancer histology images with deep neural networks / D. Ciresan//MICCAI. -2013. -№ 3.
5. Mao, Y. Iteratively training classifiers for circulating tumor cell detection / Y. Mao//Biomedical Imaging (ISBI). -2015. -№ 7.

КОМПОНЕНТ ВИЗУАЛИЗАЦИИ ГРАФОПОДОБНЫХ ДИАГРАММ НА ОСНОВЕ ИХ ФОРМАЛЬНОГО ОПИСАНИЯ В БАЗЕ ЗНАНИЙ

Бобков А. В.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: bobkov.andrey.v@gmail.com

В данной работе описан компонент визуализации графоподобных диаграмм, разработанного в рамках универсального подхода к визуализации различных знаний.

ВВЕДЕНИЕ

Пользовательский интерфейс компьютерной системы является одним из важнейших средств донесения информации конечному потребителю. Различные информационные конструкции требуют различных способов и подходов к их отображению. В настоящее время для каждого вида отображаемой информации, как правило, разрабатывается отдельный компонент визуализации. Однако такой подход трудоемок и неэффективен, поскольку добавление нового варианта отображения требует разработки нового компонента с нуля. На сегодняшний день невозможно точно классифицировать методы визуализации знаний, так как всегда можно найти новый метод, не вошедший в список [1]. Таким образом, всегда остается проблема расширяемости средств визуализации знаний.

Большинство используемых методов визуализации имеют графоподобную структуру и могут быть визуализированы с помощью графа [2]. Графом называется пара (V, E) , где E - произвольное подмножество множества $V^{(2)}$. Элементы множества V называются вершинами графа, а элементы множества E - ребрами. В силу того, что рисунок графа достаточно нагляден, а алгоритмы на графах обладают большой скоростью, они стали популярны в информатике.

Графоподобной диаграммой будем называть диаграмму, внешний вид которой можно свести к узлам и ребрам графа. Пример графоподобной диаграммы (в нотации PFC [3]) можно увидеть на рисунке 1.

Создание универсальных средств визуализации представляет собой очень сложную задачу, в связи с чем в данной работе предлагается более частное решение, позволяющее визуализировать графоподобные модели. Разработанный компонент может быть использован в любой системе, построенной с помощью технологии OSTIS [4]. OSTIS - это открытая технология для проектирования интеллектуальных компьютерных систем на основе семантических сетей с теоретико-множественной интерпретацией. Базовым языком этой технологии является SC-код. SC-код - язык внутреннего смысло-

вого представления знаний. Чтобы не работать напрямую с универсальными вариантами внешнего представления SC-кода, такими как SCn и SCg, используют специализированные компоненты пользовательского интерфейса, способные наиболее доступным образом отобразить знания конечному пользователю системы.

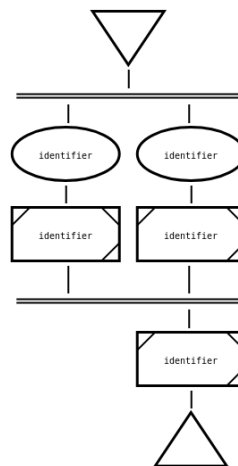


Рис. 1 – Пример графоподобной диаграммы

I. ИСПОЛЬЗУЕМЫЙ ПОДХОД

В основе разработанного компонента визуализации знаний лежит подход, основанный на разделении описания алфавита языка, его синтаксиса, и правил преобразования внутренних семантических структур в структуры на внешнем языке. Знания, хранящиеся в базе знаний системы (записанные в виде семантической модели), с помощью определенных правил перевода, преобразуются из семантической модели в синтаксическую модель [5]. Синтаксическая модель показывает, какие примитивы необходимо отобразить и каким образом они связаны между собой. При использовании такого подхода разработчику компонентов интерфейса не нужно вникать в особенности внутреннего представления знаний в системе, так как все эти знания могут быть визуализированы с помощью соответствующих синтаксических моделей. Пример синтаксической модели диаграммы 1 на языке SCg изображен на рисунке 2.

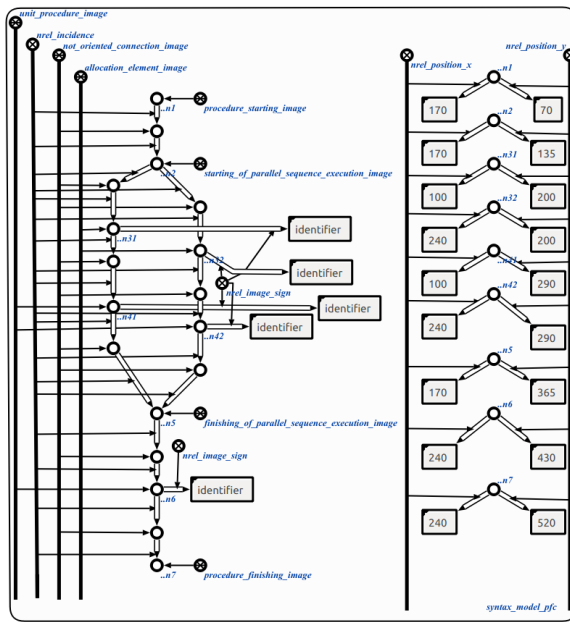


Рис. 2 – Синтаксическая модель диаграммы 1 на языке SCg

II. ОСОБЕННОСТИ СИНТАКСИЧЕСКОЙ МОДЕЛИ

Так как граф это пара, в базе знаний было выделено два множества графических примитивов, а именно узлы и ребра:

1. `element_class_image`;
2. `connector_class_image`.

Такое разделение обусловлено тем, что элемент множества `element_class_image` является независимым элементом при визуализации и не зависит от других элементов. Объекты, принадлежащие множеству `connector_class_image`, не могут существовать сами по себе, так как для их существования на схеме нужно два элемента из множества `element_class_image`.

Каждый элемент множества `element_class_image` и `connector_class_image` должен иметь свое уникальное отображение. Задать это отображение можно с помощью определенного шаблона. Такой шаблон можно описать sc-ссылкой (знаком файла, внешнего по отношению к SC-коду) и связать его с обозначаемым примитивом при помощи соответствующего отношения.

Разработанный компонент предназначен для работы в web-приложениях и работает с HTML шаблонами (задаются отношением `nrel_html_template`). В шаблоне может использоваться конструкция вида `"#{nrel_*}"`, во время визуализации эта конструкция будет заменена на значение sc-ссылки, которая связана отношением `nrel_*` с узлом, обозначающим конкретное изображение. Фрагмент базы знаний с элементом множества `element_class_image` связанным с sc-ссылкой посредством отношения `nrel_html_template` на языке SCg отображен на рисунке 3. Данный фрагмент служит для описа-

ния первого элемента диаграммы 1 (треугольника с основанием вверх).

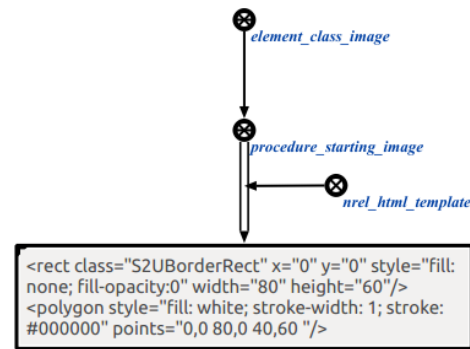


Рис. 3 – Фрагмент базы знаний с отношением `nrel_html_template` на языке SCg

Для хранения расположения элементов используется декартова система координат на плоскости. Для обозначения координат X и Y используются такой же подход, как и с шаблонами элементов (отношения `nrel_position_x` и `nrel_position_y`).

III. ЗАКЛЮЧЕНИЕ

Для решения проблем, возникающих при использовании существующих подходов визуализации знаний, был использован новый подход визуализации знаний, основанный на использовании описания синтаксической модели знаний и специализированного языка представления знаний в базе знаний системы.

Согласно этому подходу, был реализован компонент визуализации баз знаний для систем, построенных по технологии OSTIS. Разработанный компонент позволяет отображать знания с помощью любой модели визуализации, отображение которой можно свести к графоподобной диаграмме, имея только формальное описание этой модели визуализации. Разработанный компонент может быть использован в любой предметной области, для которой актуально использование моделей визуализации, которые могут быть сведены к графоподобным диаграммам.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Гаврилова, Т. А. Инженерия знаний. Модели и методы / Т. А. Гаврилова, Д. В. Кудрявцев, Д. И. Муромцев. — Лань, 2016. — 324 С.
2. Бобков, А.В. Специализированные средства визуализации баз знаний в интеллектуальных системах: дипломный проект / А. В. Бобков. - Минск : БГУИР, 2018. - 73 С.
3. Batch Control Part 2: Data Structures and Guidelines for Languages: ANSI/ISA-88.00.02-2001, 2001. – 124 с
4. Метасистема IMS.OSTIS [Электронный ресурс]. Минск, 2018. – Режим доступа: <http://ims.ostis.net/>. – Дата доступа: 1.09.2018.
5. Бобков, А. В. Семантические средства визуализации различных видов графической информации / А. В. Бобков, Бруцкий А. А, Жук А. А // ИТС 2017: материалы международной научной конференции, БГУИР, 2017 - С. 136-137.

ПРЕДСКАЗАТЕЛЬНАЯ МОДЕЛЬ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ УЯЗВИМОСТЕЙ КОМПЬЮТЕРНЫХ СИСТЕМ

Доронин А. К., Липницкий В. А.

Кафедра защиты информации, Белорусский государственный университет информатики и радиоэлектроники

Кафедра высшей математики, Военная академия Республики Беларусь

Минск, Республика Беларусь

E-mail: alixei.doronin@gmail.com, valipnitski@yandex.by

В докладе рассматривается использование конволюционных нейронных сетей в сочетании с алгоритмом представления слов в многомерном векторном пространстве GloVe для задачи предсказания критичности уязвимости, основываясь лишь на её текстовом описании. В разделе I характеризуется стандарт оценок уязвимостей CVSS. В разделе II приводится анализ распределения оценок уязвимостей из базы данных NVD. В разделе III описывается алгоритм сопоставления слов векторам (GloVe). Раздел IV содержит информацию о практических особенностях построения предсказательной модели.

ВВЕДЕНИЕ

В мире с каждым днём появляются всё новые уязвимости компьютерных систем разной степени критичности. Некоторые из них практически не представляют опасности, а другие при эксплуатации способны нанести огромный ущерб не только инфраструктуре, но и человеку. Чем быстрее будет определена степень риска недавно появившейся уязвимости, тем скорее организации смогут предпринять меры для её нейтрализации. Американская национальная база данных уязвимостей NVD содержит около 100 000 записей о различных уязвимостях компьютерных систем, найденных почти за 20 лет развития компьютерных технологий [1]. В этой базе данных (БД) присутствуют следующие поля для описания каждой уязвимости: имя уязвимости, текстовое описание, версия уязвимого ПО, экспертные оценки о критичности. Записи из данной БД представляют собой обширное экспериментальное поле для построения различных предсказательных моделей, особенно в сочетании с применением современных алгоритмов машинного обучения и обработки естественного текста. В данной работе рассматривается попытка применить метод векторного представления слов GloVe [2] в комбинации с конволюционной нейронной сетью в целях предсказания оценки (классификации) уязвимостей по их текстовому описанию.

I. АМЕРИКАНСКАЯ НАЦИОНАЛЬНАЯ БАЗА ДАННЫХ УЯЗВИМОСТЕЙ NVD

Американская Национальная база данных уязвимостей NVD (National Vulnerabilities Database) основана на списке уязвимостей из проекта CVE, запущенном MITRE в 1999 году. CVE (Common Vulnerabilities and Exposures) - это общедоступный список записей, каждый из которых содержит идентификационный номер, описание и, по крайней мере, одну общедо-

ступную ссылку для каждой уязвимости. Записи из списка CVE используются в многочисленных продуктах и услугах по кибербезопасности со всего мира, в том числе и NVD. Это означает, что база данных уязвимостей NVD полностью синхронизирована с CVE, однако содержит некоторую дополнительную информацию и предоставляет расширенные функции поиска. Все записи в базе данных NVD имеют 14 полей, из них основными являются следующие:

1. CVE Id: уникальный ID (номер) уязвимости (например, CVE-2011-1585)
2. Date published: дата первой публикации о данной уязвимости
3. Date modified: последняя дата изменения записи
4. Summary: текстовое описание уязвимости

Оставшиеся поля относятся к параметрам оценки уязвимости, выставляемой в соответствии со стандартом CVSS [3]:

1. CVSS Base: базовая оценка уязвимости (десятичное значение от 0 до 10, например, 9.5)
2. CVSS Impact: оценка элементов воздействия (десятичное значение от 0 до 10, например, 9.5)
3. CVSS Exploit: оценка возможности эксплуатации (десятичное значение от 0 до 10, например, 9.5)
4. CVSS Access vector (AV): вектор доступа (локальный/сетевой/локально-сетевой)
5. CVSS Access complexity (AC): сложность доступа (низкая/средняя/высокая)
6. CVSS Authentication (Au): уровень требуемой аутентификации (нулевая/однократная/многократная)
7. CVSS Confidentiality impact (C): воздействие на конфиденциальность (полное/частичное/нулевое)

8. CVSS Integrity impact (I): воздействие на целостность (полное/частичное/нулевое)
9. CVSS Availability impact (A): воздействие на доступность (полное/частичное/нулевое)
10. CVSS Vector: базовый вектор уязвимости содержит в себе значения других полей (например, AV:N/AC:M/Au:N/C:N/I:P/A:C)

II. АНАЛИЗ БАЗЫ ДАННЫХ NVD

Для проведения дальнейших исследований критерий оценок CVSS Base был выбран в качестве «главной оценки» уязвимости (рисунок 1).

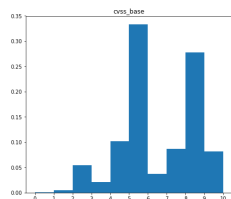


Рис. 1 – Гистограмма распределения оценок CVSS Base

На гистограмме видно, что примерно половина всех оценок сгруппирована вокруг отметки в 4.5 балла, а остальная половина - вокруг 8. Напомним, что данные оценки выставляются экспертами. Перед нами поставлена задача смоделировать систему, которая смогла бы автоматически предсказывать, к какой группе относится новая уязвимость: к группе с низкими оценками (меньше или равно 5 баллам), или же к группе с высокими оценками (более 5 баллов). В качестве входных данных для данной модели наиболее интересным выглядит текстовое описание уязвимости. Таким образом, имеем задачу классификации с двумя классами: класс 0 - уязвимости с низкими оценками, класс 1 - с высокими оценками.

III. ВЕКТОРНОЕ ПРЕДСТАВЛЕНИЕ СЛОВ

Очевидно, что для построения предсказательной модели по текстовому описанию необходимо каким-либо образом осуществить переход от слов к некоторому числовому их эквиваленту. Одним из решений данной задачи является алгоритм GloVe, разработанный в 2013 году [2]. Нами использовался наиболее простой словарь GloVe, содержащий в себе пары "слово – вектор": размерность каждого вектора - 50, общий объём словаря - 400.000. Применение данного словаря позволяет сделать переход к 50-мерному векторному пространству, в котором каждая точка - это вектор, поставленный в соответствие некоторому слову. Ключевая особенность данного алгоритма в том, что на этапе построения вектора слова учитывается контекст окружающих его слов. Это позволяет говорить о том, что при близком расположении точек в векторном пространстве значения соответствующих слов с высокой вероятностью будут так же близки по смыслу.

IV. ПОСТРОЕНИЕ И ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

В качестве предиктивного алгоритма было принято решение использовать конволюционную нейронную сеть, так как данный тип нейросетей обычно показывает хорошие результаты в задачах обработки естественного текста. В результате была построена многослойная нейронная сеть с 2 выходами, классифицирующая уязвимости по их текстовому описанию с точностью 85% на отложенной тестовой выборке. На вход нейросети подаётся текстовое описание уязвимости, которое преобразуется в вектор в соответствии со словарём GloVe. Далее преобразованные данные проходят три конволюционных слоя (conv1d), после каждого из них - следует обобщающий (MaxPooling) слой. С целью уменьшения переобучения затем идут три полносвязных (Dense) слоя в сочетании с исключаящими (Dropout) слоями. Нейросеть обучалась на 20% от всей выборки, из которых 20% были использованы в качестве валидирующего набора данных. Более подробно с исходным кодом и комментариями можно ознакомиться по ссылке: https://github.com/teacherlex/cve_vulns_classifier

ЗАКЛЮЧЕНИЕ

В докладе дано построение модели, предсказывающей критичность той или иной уязвимости, основываясь лишь на её текстовом описании. В качестве звена, преобразующего текст в векторное пространство, используется словарь проекта GloVe, а в качестве основного обучающегося алгоритма - многослойная конволюционная нейронная сеть. Наилучшая из построенных нами моделей достигла показателя 91% по метрике площади под AUC-ROC кривой. Отметим, что построенную модель можно использовать для оценки любого текста, а не только являющегося непосредственно описанием уязвимости. Например, используя данную модель, можно построить фильтр, отсеивающий лишь наиболее «критичные» тексты. Кроме того, перспективным может оказаться переход к задаче регрессии: в этом случае можно отсортировать тексты в порядке критичности содержащейся в них информации с точки зрения уязвимостей компьютерных систем.

СПИСОК ЛИТЕРАТУРЫ

1. National Vulnerability Database [Электронный ресурс] / The Natural Language Processing Group at Stanford University. — Режим доступа: <https://cve.mitre.org>. — Дата доступа: 25.06.2018.
2. GloVe: Global Vectors for Word Representation [Электронный ресурс] / The Natural Language Processing Group at Stanford University. — Режим доступа: <https://nlp.stanford.edu/projects/glove/>. — Дата доступа: 07.07.2018.
3. Общая система оценки уязвимостей CVSS [Электронный ресурс] / BIS Expert. — Режим доступа: <http://bis-expert.ru/blog/5345/43124>. — Дата доступа: 18.07.2018.

СРЕДСТВА ВЕРИФИКАЦИИ МАШИН ОБРАБОТКИ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Дюбина Е. А., Давыденко И. Т.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: dziubina.el@gmail.com, davydenko@bsuir.by

Рассматриваются средства верификации машин обработки знаний интеллектуальных систем, построенных по Технологии OSTIS.

ВВЕДЕНИЕ

Важным этапом тестирования программного продукта является верификация. На этом этапе определяется, удовлетворяют ли результаты текущего этапа разработки требованиям, установленным в начале этого этапа. Верификация осуществляется самими разработчиками и позволяет убедиться, что продукт разработан «правильно».

Верификация позволяет выявлять дефекты на ранних этапах разработки, а следовательно, затраты на исправление обнаруженных дефектов минимальны.

Интеллектуальные системы, как и другие программные продукты, нуждаются в проведении верификации для повышения их качества и сокращения времени на выявление дефектов.

I. МАШИНА ОБРАБОТКИ ЗНАНИЙ

Машина обработки знаний (МОЗ) является важнейшей частью интеллектуальных систем, так как именно её возможностями определяется функционал системы в целом, возможность давать ответы на нетривиальные вопросы пользователя и способность решать различные задачи [1].

В основе данной работы лежит рассмотрение систем, построенных по Технологии OSTIS. Такие системы названы ostis-системами [2]. В качестве модели представления знаний в ostis-системах используется семантическая сеть с теоретико-множественной интерпретацией.

МОЗ, построенная по Технологии OSTIS, представляет собой графодинамическую машину, включающую в себя систему агентов и общую память. Отличительной особенностью таких МОЗ является то, что агенты взаимодействуют друг с другом исключительно через общую память.

II. ВЕРИФИКАЦИЯ МАШИНЫ ОБРАБОТКИ ЗНАНИЙ

На сегодняшний день проверка корректности работы МОЗ осуществляется путём ручной верификации. Это долгий и крайне трудоёмкий процесс. Осуществляемая верификация относится к интеграционному типу (верификации взаи-

модействия МОЗ и пользовательского интерфейса) и не позволяет в полной мере определить, правильно ли работает МОЗ как отдельный компонент интеллектуальной системы. При этом не наблюдается снижения затрат при многократном проведении верификации.

Верификация МОЗ — это сложный процесс, в рамках которого решаются следующие задачи:

1. Определить основную цель верификации и задачи, выполнение которых необходимо для достижения цели.
2. Установить критерии завершения верификации для данного этапа разработки.
3. Осуществить анализ предметной области и разработать тест-кейсы.
4. Запустить верифицируемый компонент на исполнение с заданными аргументами из тест-кейса.
5. Сравнить ожидаемый результат с полученным.
6. Сформировать отчёт с результатами верификации.
7. Проанализировать отчёт.
8. Разработать план дальнейших действий на основе полученных результатов.

В зависимости от сложности разрабатываемой МОЗ, её верификацией может заниматься как один человек, так и целая команда специалистов. Полностью исключить человека из процесса верификации невозможно. Тем не менее можно автоматизировать некоторые её этапы (5-7).

Преимущества автоматизированной верификации:

- минимизация затрат при многократном проведении;
- скорость выполнения тестов может намного превосходить возможности человека;
- отсутствие влияния человеческого фактора, например, усталости, невнимательности и так далее;
- способность средств автоматизации выполнить проверки, непосильные для человека в силу своей сложности, скорости или иных факторов;
- способность средств автоматизации собирать, сохранять, агрегировать и представ-

лять в удобной для восприятия человеком форме большие объёмы информации.

Недостатки автоматизированной верификации:

- затраты на разработку и сопровождение тест-кейсов;
- в случае кардинального изменения требований появляется необходимость в изменении ранее разработанных тестов.

Принимая во внимание критичную значимость МОЗ для интеллектуальной системы, описанные выше достоинства перевешивают недостатки.

Верификация осуществляется на модульном (компонентном) уровне детализации, где объектами верификации являются небольшие части приложения. Важной особенностью верификации на модульном уровне является проверка корректности работы компонента в изолированной среде. При использовании объектно-ориентированной парадигмы программирования объектами верификации являются методы классов, сами классы, взаимодействие классов. Также верификации могут подлежать компоненты более высокого уровня абстракции: небольшие библиотеки и отдельные части системы. В рамках МОЗ объектами верификации являются агенты и процедуры, используемые в агентах.

Как правило, разработчики программного обеспечения используют специальные технологии и инструментальные средства автоматизированной верификации, что позволяет существенно ускорить и упростить процесс верификации.

III. СРЕДСТВА ВЕРИФИКАЦИИ МАШИН ОБРАБОТКИ ЗНАНИЙ

Средства верификации МОЗ должны удовлетворять следующим требованиям:

- осуществлять автоматизированную верификацию на основе тест-кейсов;
- осуществлять верификацию компонентов МОЗ в изолированной среде;
- осуществлять валидацию тест-кейсов;
- по окончании верификации формировать отчёт с результатами верификации;
- тест-кейсы и отчёты о проведённой верификации должны являться частью базы знаний.

В рамках предлагаемого подхода средства автоматизированной верификации МОЗ представляют собой самостоятельную программную систему, реализованную средствами Технологии OSTIS, то есть имеющую:

- базу знаний, содержащую языковые средства описания тест-кейсов, этапов верификации, отчётов;

- МОЗ, состоящую из коллектива агентов, обеспечивающих автоматизированную верификацию компонентов МОЗ, а также генерацию и редактирование тест-кейсов;
- пользовательский интерфейс.

Средства автоматизированной верификации позволяют автоматизировать процесс запуска верифицируемого компонента на исполнение, оценки успешности проверки тест-кейса и формирования отчёта, а также частично автоматизировать механизмы управления тест-кейсами.

Так как тест-кейсы являются частью базы знаний и обладают высокой степенью формализованности, то они могут быть полезными не только для разработчиков, но и для конечных пользователей, которые могут использовать их в качестве наглядной инструкции о том, с какими аргументами можно вызывать компонент МОЗ и какой результат при этом ожидать.

Каждый тест-кейс имеет соответствующую спецификацию. В спецификацию тест-кейса входят:

- основной идентификатор;
- принадлежность к тест-набору;
- указание объекта верификации;
- формальное описание ожидаемого результата;
- формальное описание фактического результата;
- формальное описание аргументов;
- формальное описание контекста действия (некоторого фрагмента базы знаний, в рамках которого осуществляется работа объекта верификации);
- строгое, полное, однозначно понимаемое описание назначения тест-кейса, условий верификации, ожидаемого результата на естественном языке;
- указание теста (для автоматизированных тест-кейсов).

Таким образом, использование средств автоматизированной верификации МОЗ позволяет существенно ускорить, упростить и унифицировать процесс верификации. Являясь частью библиотеки многократно используемых компонентов, данные средства могут быть легко интегрированы в любую интеллектуальную систему, построенную по Технологии OSTIS.

1. Шункевич, Д. В. Модели и средства компонентного проектирования машин обработки знаний на основе семантических сетей / Д. В. Шункевич // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2013): материалы III Междунар. научн. -техн. конф (Минск, 21-23 февраля 2013г.). – Минск: БГУИР, 2013. – С. 269-280.
2. База знаний IMS // Метасистема IMS [Электронный ресурс]. – 2017. – Режим доступа: <http://ims.ostis.net>. – Дата доступа: 10.09.2018.

ОНТОЛОГИЯ ПРЕДМЕТНОЙ ОБЛАСТИ ПОДДЕРЖКИ, ПОДГОТОВКИ И ПРОВЕДЕНИЯ КОНФЕРЕНЦИИ

Астахова Н. М., Федорович А. Г.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {astnadezda07, artemfedorovi4}@gmail.com

В данной статье рассмотрена структура онтологии предметной области поддержки, подготовки и проведения конференций

ВВЕДЕНИЕ

При подготовке и проведении конференций возникает ряд прикладных задач, связанных с организацией информации об участниках конференции, установлением очередности проведения докладов, а также проверкой корректности оформления поданных заявок и докладов. При решении данных задач возникает необходимость в систематизации данного объема информации, для успешной в дальнейшем организации конференции любого уровня и направленности, что и подразумевается под поддержкой, подготовкой и проведением конференции.

Далее будет рассмотрена онтология[1] предметной области поддержки, подготовки и проведения конференций, которая может быть использована для спецификации внешних языков представления знаний[2], таких как доклады, статьи и т.п.

I. ОПИСАНИЕ ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ

Основополагающим понятием Предметной области поддержки, подготовки и проведения конференций является *конференция*.

Конференция - собрание организаций и лиц для обсуждения, решения и представления некоторых вопросов. Конференция бывает нескольких видов:

- В зависимости от направления:
 - Научная конференция
 - Бизнес конференция
 - Пресс-конференция
- В зависимости от охваченной территории:
 - Локальные
 - Областные
 - Республиканские
 - Международные
- В зависимости от формы проведения:
 - Очная
 - Заочная
 - Веб-конференция[3]
 - Видеоконференция

Научная конференция - форма организации научной деятельности, при которой исследователи представляют и обсуждают свои рабо-

ты. На данный момент выделено три подвида научной конференции:

1. **Научно-теоретическая.** Обсуждение новых разработок, исследований, открытий, изучение статистических данных.
2. **Научно-практическая.** Выступающие строят свои доклады, основываясь на личных наблюдениях, исследованиях по обсуждаемой теме или на исследованиях своих коллег.
3. **Научно-техническая.** Это своего рода деловое общение студентов и преподавателей с представителями компаний.

Бизнес-конференция - мероприятие, на котором обсуждаются бизнес-проблемы, особенности действующего законодательства, политика государства по отношению к определенным отраслям.

Пресс-конференция - встреча политических, общественных, научных деятелей с представителями печати, радио, телевидения для беседы по вопросам, представляющим большой общественный интерес.

Участник - абсолютное понятие, которое обозначает лицо, которое принимает участие в каком-либо мероприятии.

Докладчик* - отношение, первой компонентой связок является знак *конференции*, на которую *участник*, знак участника является второй компонентой связок, предоставляет свой *доклад*.

Слушатель* - отношение, первой компонентой связок которого является знак *конференции*, на которой присутствует *участник*, знак которого является второй компонентой связок.

Организационный комитет* - отношение, которое связывает знак *конференции* и знак *участника*, занимающегося организацией *конференции*.

Техническая и информационная поддержка конференции* - отношение, которое связывает знак *конференции* и знак *участника*, занимающегося организацией *конференции*.

Доклад - это абсолютное понятие, которое значит развернутое устное сообщение на какую-либо тему, сделанное публично. Каждый доклад может быть описан такими отношениями как:

- *название доклада**
- *язык доклада**
- *дата выступления**
- *время выступления**
- *опубликованный материал**
- *адрес**
- *докладчик**
- *автор**

Название доклада* - отношение, первой компонентой связок которого является знак *доклада*, а второй - текст, являющийся названием данного *доклада*.

Язык доклада* - отношение, первой компонентой связок которого является знак *доклада*, а второй - знак языка, на котором представлен данный *доклад*.

Дата выступления* - отношение, первой компонентой связок которого является знак *доклада*, а второй - дата представления *доклада*.

Время выступления* - отношение, первой компонентой связок которого является знак *доклада*, а второй - время проведения.

Опубликованный материал* - отношение, связки которого связывают между собой некоторую сущность и субъект, который является *статьей* данной сущности. Это отношение показывает изданный материал, который хранится в электронном или печатном виде.

Адрес* - отношение, первой компонентой связок которого является знак *доклада*, а второй - знак *места проведения* конференции.

Автор* - отношение, связки которого связывают между собой некоторую сущность и субъект, который является *автором* (создателем) данной сущности.

Для описания статьи используются такие связки как:

- *название статьи**
- *язык доклада**
- *аннотация**
- *автор**
- *год публикации**

Далее рассмотрим пример формализованной статьи на примере данной статьи: **статья_№1**

=> *название публикации**:

- *Онтология предметной области поддержки подготовки и проведения конференций.*

=> *язык публикации**:

- *русский язык*

=> *аннотация**:

- *В данной статье рассмотрена структура онтологии предметной области поддержки, подготовки и проведении конференций.*

=> *автор**:

- *Федорович А.Г.*
- *Астахова Н.М.*

=> *год выпуска**:

- *2018*

Далее будет представлен *доклад* на формальном языке:

Доклад 3.08 OSTIS-2016 Модели предметных областей в системах поддержки принятия решений на основе мониторинга информационного пространства => *название доклада**:

- *Модели предметных областей в системах поддержки принятия решений на основе мониторинга информационного пространства.*

=> *язык доклада**:

- *русский язык*

=> *автор**:

- *Додонов А.Г.*
- *Ландэ Д.В.*
- *Коваленко Т.В.*

=> *докладчик**:

- *Додонов А.Г.*

=> *опубликованный материал**:

- *Модели предметных областей в системах поддержки принятия решений на основе мониторинга информационного пространства.*

=> *дата выступления**:

- *16 февраля 2016*

=> *время выступления**:

- *17:30 - 18:30*

=> *адрес**:

- *Республика Беларусь, г. Минск, ул. Платонова, д. 39, ауд. 414*

II. ЗАКЛЮЧЕНИЕ

В работе приведена структура онтологии предметной области поддержки, подготовки и проведения конференции, используемая для решения задач систематизации и организации информации о конференции.

1. Davydenko, I. T. Ontology-based knowledgebase design, / I. Davydenko // Open Semantic Technologies for Intelligent Systems: материалы междунар. науч.-техн. конф./ редкол.: В. В. Голенков (отв. ред.) [и др.], ISSN 2415-7740; Вып.1 (Минск, 16-18 февр. 2017г.). – Минск: БГУИР, 2017. – С. 57–72.
2. Борискин, А. С., Жуков, И. И., Корончик, Д. Н., Садовский, М. Е., Хусаинов, А. Ф. - Онтологическое проектирование пользовательских интерфейсов интеллектуальных систем. -- В книге Междунар. научн.-техн. конф. . «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS –2017). Материалы конф. -- Минск: БГУИР, 2017
3. Алексеева Т. Б., Гладкая И. В., Сеницына А. И. Интернет-конференция как формаформа научной коммуникации молодых исследователей // Современные проблемы науки и образования. — 2013. — № 5.

АРХИТЕКТУРА СОВРЕМЕННОЙ НЕЙРОННОЙ СЕТИ ДЛЯ СЕГМЕНТАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Гончаров И. В., Потешкин А. С., Курулев А. П.

Кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {ivan98goncharov, artem.poteshkin92}@gmail.com

В статье излагается исследование одной из современной нейронной сети для сегментации объектов на изображениях на примере Mask R-CNN, разработанной Facebook Research.

ВВЕДЕНИЕ

Возможность видеть и распознавать объекты – естественная и привычная возможность для человека. Однако для компьютера пока что – это чрезвычайно сложная задача. На данный момент нейронные сети способны выполнять куда более сложные и интересные задания по обработке изображений. А именно, сеть с архитектурой Mask R-CNN позволяет выделять на фотографиях контуры («маски») экземпляров разных объектов, даже если таких экземпляров несколько, они имеют различный размер и частично перекрываются. Сеть так же способна к распознаванию поз людей.

I. Типы задач компьютерного зрения

Обычно современные задачи компьютерного зрения разделяют на четыре вида :

- Classification – классификация изображения по типу объекта, которое оно содержит.
- Semantic segmentation – определение всех пикселей объектов определённого класса или фона на изображении. Если несколько объектов одного класса перекрываются, их пиксели никак не отделяются друг от друга.
- Object detection – обнаружение всех объектов указанных классов и определение охватывающей рамки для каждого из них.
- Instance segmentation – определение пикселей, принадлежащих каждому объекту каждого класса по отдельности.

На примере изображения с воздушными шарами из (рис. 1) это можно проиллюстрировать так:

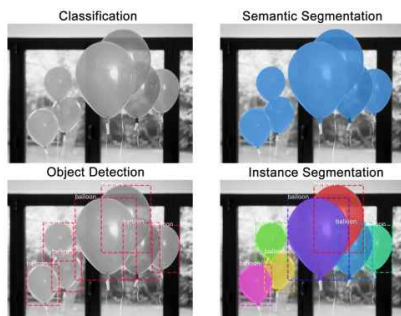


Рис. 1 – Типы задач компьютерного зрения

II. КОНЦЕПЦИИ, ЛЕЖАЩИЕ В ОСНОВЕ В MASK R-CNN

Концепции, лежащие в основе в Mask R-CNN прошли поэтапное развитие через архитектуры нескольких промежуточных нейросетей, решавших разные задачи из приведённого выше списка.

Convolution Neural Networks состоит из сверточных слоев, которые позволяют объединять значения расположенных рядом пикселей и выделять более обобщённые признаки изображения. Для этого по картинке последовательно скользят квадратным окном небольшого размера (3x3, 5x5 пикселей и т.п.) называемым ядром. Каждый элемент ядра имеет свой весовой коэффициент, умножаемый на значение того пикселя изображения, на который в данный момент наложен элемент ядра. Затем полученные для всего окна числа складываются, и эта взвешенная сумма даёт значение очередного признака. Для получения матрицы («карты») признаков всего изображения, ядро последовательно сдвигается по горизонтали и вертикали. В следующих слоях операция свертки применяется уже к картам признаков, полученным из предыдущих слоёв (рис.2).

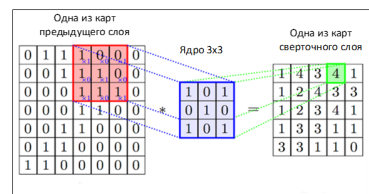


Рис. 2 – Графический процесс получения «карты» признаков изображения

Архитектура сети R-CNN (Regions With CNNs) была разработана командой из UC Berkley для применения Convolution Neural Networks к задаче object detection. Процедуру детектирования объектов сетью R-CNN можно разделить на следующие шаги:

- Выделение регионов-кандидатов при помощи Selective Search.
- Преобразование региона в размер, принимаемый CNN CaffeNet.

- Получение при помощи CNN 4096-размерного вектора признаков.
- Проведение N бинарных классификаций каждого вектора признаков при помощи N линейных SVM.
- Линейная регрессия параметров рамки региона для более точного охвата объекта.

Fast R-CNN предложили ускорить процесс R-CNN за счёт пары модификаций: пропускать через CNN не каждый из 2000 регионов-кандидатов по отдельности, а всё изображение целиком. Предложенные регионы потом накладываются на полученную общую карту признаков. Преобразование признаков, попавших в разные регионы, к фиксированному размеру производилось при помощи процедуры RoIPooling. Окно региона шириной w и высотой h делилось на сетку, имеющую HW ячеек размером $h/H \times w/W$. По каждой такой ячейке проводился Max Pooling для выбора только одного значения, давая результирующую матрицу признаков HW .

После улучшений, сделанных в Fast R-CNN, самым узким местом нейросети оказался механизм генерации регионов-кандидатов. В 2015 команда из Microsoft Research смогла сделать этот этап значительно более быстрым. Они предложили вычислять регионы не по изначальному изображению, а по карте признаков, полученных из CNN (рис.3). Для этого был добавлен модуль Region Proposal Network (RPN). В рамках RPN по извлечённым CNN признакам скользят «мини-нейросеть» с 3×3 окном. Полученные с её помощью значения передаются в два параллельных полносвязанных слоя: box-regression layer (reg) и box-classification layer (cls).

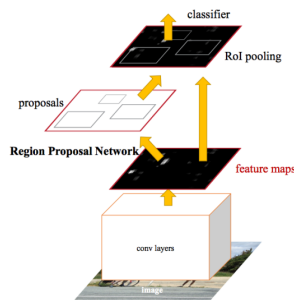


Рис. 3 – Архитектура Faster R-CNN

III. MASK R-CNN

Mask R-CNN развивает архитектуру Faster R-CNN путём добавления ещё одной ветки, которая предсказывает положение маски, покрывающей найденный объект, и решает уже задачу instance segmentation (рис.4).

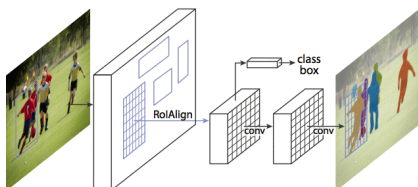


Рис. 4 – Архитектура Mask R-CNN

Маска представляет собой просто прямоугольную матрицу, в которой 1 на некоторой позиции означает принадлежность соответствующего пикселя объекту заданного класса, 0 — что пиксель объекту не принадлежит. Выделение маски происходит в class-agnostic стиле: маски предсказываются отдельно для каждого класса, без предварительного знания, что изображено в регионе, и потом просто выбирается маска класса, победившего в независимом классификаторе. Утверждается, что такой подход более эффективен. Одна из основных модификаций, возникших из-за необходимости предсказывать маску — изменение процедуры RoIPool (вычисляющей матрицу признаков для региона-кандидата) на так называемую RoIAlign. Дело в том, что карта признаков, полученная из CNN, имеет меньший размер, чем исходное изображение, и регион, охватывающий на изображении целочисленное количество пикселей, не получается отобразить в пропорциональный регион карты с целочисленным количеством признаков (рис.5).

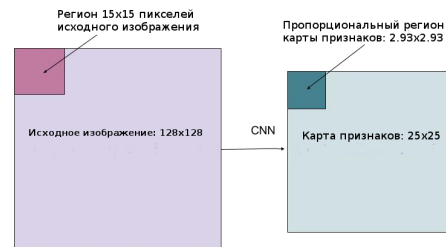


Рис. 5 – Работа процедуры RoIAlign

RoIPool проблема решалась просто округлением дробных значений до целых. Такой подход нормально работает при выделении охватывающей рамки, но вычисленная на основе таких данных маска получается слишком неточной. В противоположность этому, в RoIAlign не используется округление, все числа остаются действительными, а для вычисления значений признаков используется билинейная интерполяция по четырём ближайшим целочисленным точкам.

IV. ЗАКЛЮЧЕНИЕ

Сегментация объектов - одна из перспективных сфер для развития нейронных сетей. Данную технологию можно использовать на дорогах, в аэропортах, на улицах и т.д. Mask R-CNN стала новым уровнем искусства с точки зрения сегментации экземпляров.

СПИСОК ЛИТЕРАТУРЫ

1. R. Beith: Computer Vision: Algorithms and Applications, -2010. -979c
2. <https://blog.athelas.com/a-brief-history-of-cnns-in-image-segmentation-from-r-cnn-to-mask-r-cnn-34ea83205de4>
3. <https://engineering.matterport.com/splash-of-color-instance-segmentation-with-mask-r-cnn-and-tensorflow-7c761e238b46>

АКТУАЛЬНЫЕ ПРОБЛЕМЫ В ОБЛАСТИ РАЗРАБОТКИ ГИБРИДНЫХ БАЗ ЗНАНИЙ

Гулякина Н.А., Давыденко И.Т., Зверуго А.В.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: guliakina@bsuir.by, davydenko@bsuir.by

В работе рассмотрены основные актуальные проблемы в области разработки гибридных баз знаний. Предложен подход к их решению, основанный на использовании Технологии OSTIS

ВВЕДЕНИЕ

В настоящее время системы, основанные на знаниях (СОЗ), находят применение в различных сферах человеческой деятельности. Среди них выделяются медицина, образование, справочные системы и многие другие. Согласно маркетинговым исследованиям доля систем, основанных на знаниях, на рынке «умных машин» составит к 2019 году около 46%. При этом в настоящее время объем рынка таких систем увеличивается каждый год на 15% [1].

Ключевым компонентом в системах, основанных на знаниях является база знаний. В данной работе база знаний определяется как систематизированная совокупность всех знаний, представленных на формальном языке и необходимых для функционирования соответствующей системы, основанной на знаниях. Таким образом, качество интеллектуальной системы во многом определяется качеством ее базы знаний.

Повышение сложности задач, решаемых компьютерными системами, основанными на знаниях, приводит к усложнению структуры баз знаний таких систем и к расширению многообразия видов знаний, входящих в состав базы знаний.

I. КОМПЛЕКСНЫЕ ЗАДАЧИ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ

Среди многообразия видов знаний выделяются факты, различные спецификации объектов, закономерности, процессы, правила и др.

При этом информация в базу знаний поступает из разных источников, кроме того, существует большое число факторов, влияющих на содержимое БЗ, таких, как изменение условий использования системы, изменения в технологиях и стандартах и др (см. рис. 1).

Расширение сфер применения СОЗ требует от таких систем способности решать комплексные задачи [3], т.е. задачи, требующие совместного использования различных видов знаний и различных моделей решения задач.

Примерами комплексных задач являются следующие: задача понимания текстов естественного языка, речевых сообщений, изображений; задача автоматизации адаптивного обуче-

ния; задача комплексной автоматизации различных предприятий и др.

В свою очередь базу знаний в рамках которой обеспечивается совместное использование различных видов знаний при решении любой комплексной задачи, будем называть *гибридной базой знаний* [3].



Рис. 1 – Источники информации, поступающей в базу знаний

II. АКТУАЛЬНЫЕ ПРОБЛЕМЫ В ОБЛАСТИ РАЗРАБОТКИ ГИБРИДНЫХ БАЗ ЗНАНИЙ

На основе анализа задач, решаемых интеллектуальными системами были сформулированы требования, предъявляемые к гибридным базам знаний, к ним относятся:

- возможность согласованного использования различных видов знаний в рамках одной и той же базы знаний, т.е. обеспечение совместимости знаний различного вида;
- наличие структуры, учитывающей различные аспекты спецификации сущностей, описываемых в базе знаний;
- возможность представления в базе знаний метазнаний различного уровня;
- удобство обработки базы знаний;
- модифицируемость базы знаний, т.е. легкость внесения изменений в такую базу знаний.

На сегодняшний день выделяются следующие актуальные проблемы в области разработки гибридных баз знаний:

- трудоемкость одновременного использования моделей представления различных видов знаний
- несовместимость уже разработанных компонентов баз знаний, что приводит к необходимости повторной разработки уже существующих решений;
- сложность внесения изменений в базу знаний;
- существующие инструментальные средства не в полной мере обеспечивают комплексную поддержку разработки, а также не обладают достаточной расширяемостью.

Как следствие разработка баз знаний – трудоемкий и продолжительный процесс, требующий высокого уровня квалификации разработчиков, что, в свою очередь, влечет за собой дефицит специалистов в данной области, а также высокую стоимость баз знаний, а, следовательно, интеллектуальных систем.

Среди попыток решения указанных проблем выделяются следующие:

- разработка интегрированных моделей и языков представления знаний [5];
- разработка онтологии верхнего уровня и библиотек прикладных онтологий [2, 6];
- переход к редактированию в общей памяти вместо редактирования исходных текстов;
- разработка новых инструментов создания баз знаний – редакторов, средств интеграции, верификации и др. [4]

Однако существующие реализации указанных подходов, как правило, направлены на решение какой-либо одной из указанных проблем и не учитывают необходимости решения всех указанных проблем в комплексе [4].

Основной причиной указанных выше проблем является отсутствие совместимости представляемых знаний, которая включает два аспекта: синтаксическая совместимость и семантическая совместимость.

Синтаксическая совместимость заключается в обеспечении унификации формы представления знаний и решается путем разработки новых моделей представления знаний и соответствующих им языков.

Семантическая совместимость подразумевает однозначную и единую для всех компонентов базы знаний интерпретацию используемых понятий, одной из попыток обеспечения которой является разработка онтологий верхнего уровня.

III. ПРЕДЛАГАЕМЫЙ ПОДХОД

Для решения проблемы совместимости представляемых знаний необходимо рассмотреть оба аспекта в совокупности, при этом тре-

буется наличие единого унифицированного формального базиса. Такой подход положен в основу данной работы.

В основе работы используется *Технология OSTIS*, предполагающая кодирование информации в виде семантических сетей с теоретико-множественной интерпретацией. Такой способ кодирования назван *SC-код* [3]. Также в рамках технологии используются алгоритмы и модели интеграции семантических сетей и обработка знаний на основе агентно-ориентированного подхода.

Основной особенностью *SC-кода* является теоретико-множественная интерпретация его элементов, что обеспечивает однозначность представления информации, т.е. *SC-код* является основой для решения проблемы совместимости представляемых знаний.

Для решения рассмотренных проблем предлагается использовать принципы: унификации представления знаний; стратификации базы знаний, т.е. построения ее иерархической структуры; структуризации базы знаний по различным признакам; описание структуры базы знаний в самой базе знаний; согласованная коллективная разработка; компонентный подход; автоматизация разработки; разработка инструментальных средств на основе *Технологии OSTIS* для обеспечения их легкой расширяемости.

СПИСОК ЛИТЕРАТУРЫ

1. “Smart Machines: Technologies and Global Markets”, BCC Research, March 2016 [Electronic resource]. - Mode of access: <https://www.bccresearch.com/market-research/instrumentation-and-sensors/smart-machines-tech-markets-report-ias094b.html>. - Date of access: 20.05.2018.
2. Борисов, А. Н. Построение интеллектуальных систем, основанных на знаниях, с повторным использованием компонентов / А. Н. Борисов // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2014) : материалы IV Междунар. техн.-конф., Минск, 20–22 февр. 2014 г. / БГУИР ; редкол.: В. В. Голенков (отв. ред.) [и др.]. — Минск, 2014. — С. 97–102.
3. Семантическая модель представления и обработки баз знаний / В.В. Голенков, Н. А. Гулякина, И. Т. Давыденко, Д. В. Шункевич // Аналитика и управление данными в областях с интенсивным использованием данных: сб. науч. тр. XIX Междунар. конф. DAMDID / RCDL'2017, Москва, 10–13 окт. 2017 г. / Федер. исслед. центр «Информатика и управление» Рос. акад. наук [и др.] ; под ред. Л. А. Калиниченко [и др.] — М., 2017. — С. 412–419.
4. Грибова, В. В. Проект IASaaS. Комплекс для интеллектуальных систем на основе облачных вычислений / В. В. Грибова // Искусств. интеллект и принятие решений. — 2011. — № 1. — С. 27–35.
5. Загорулько, Ю. А. О концепции интегрированной модели представления знаний / Ю. А. Загорулько // Изв. Том. политехн. ун-та. — 2013. — Т. 322, № 5. — С. 98–103.
6. Лапшин, В. А. Онтологии в компьютерных системах / В. А. Лапшин. — М. : Науч. мир, 2010. — 224 с.

АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПОИСКА ЭЛЕМЕНТОВ ПОЗВОНКА НА КТ-ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ АНАЛИЗА КОНТУРОВ

Губарева О. И.

Кафедра информационных технологий, Гомельский государственный университет имени П.О. Сухого
Гомель, Республика Беларусь
E-mail: olga.gubareva29@gmail.com

В данной работе рассматривается реализация поиска элементов позвонка на КТ-изображениях с помощью сегментации изображения. Данное приложение было разработано с использованием сверточной нейронной сети.

ВВЕДЕНИЕ

Диагностическая визуализация – важный инструмент в медицине. Магнитно-резонансная томография (МРТ), компьютерная томография (КТ), цифровая маммография и другие методы визуализации обеспечивают эффективное средство для неинвазивного картирования анатомии субъекта. Эти технологии значительно расширили знания о анатомии для медицинских исследований и являются важным компонентом в диагностике и планировании лечения.

В связи с растущим размером и количеством данных медицинских изображений, появилась необходимость в использовании информационных технологий для облегчения обработки и анализа. В частности, различные алгоритмы для разграничения анатомических структур и других интересующих областей становятся все более важными для оказания помощи и автоматизации конкретных медицинских задач. Эти алгоритмы, называемые алгоритмами сегментации изображений, играют жизненно важную роль в многочисленных приложениях для биомедицинских изображений, таких как количественное определение объемов ткани, диагноз, локализация патологии, исследование анатомической структуры, планирование лечения и компьютерно-интегрированная хирургия.

I. ПОИСК ЭЛЕМЕНТОВ ПОЗВОНКА

В данной работе поиск элементов позвонка осуществляется с использованием Inception-v3 модели [1]. Данная модель является одной из предварительно подготовленных моделей TensorFlow. Это переосмысление исходной структуры компьютерного зрения после Inception-v1, Inception-v2 в 2015 году. Модель Inception-v3 обучается наборам данных ImageNet, содержащим информацию, которая может идентифицировать 1000 классов в ImageNet. Для того, чтобы использовать данную модель необходимо переобучить последний слой Inception для новых категорий с использованием передачи знаний. Архитектура данной модели представлена на рисунке 1.

Передача знаний – это новый метод машинного обучения, который может использовать существующие знания, полученные из одной среды, и решать другую новую проблему, которая отличается, но имеет некоторое отношение к старой проблеме [2]. Например, мы можем применить знания, полученные от решения проблемы одного вида цветов, к изучению схожей проблемы у другого. По сравнению с традиционной нейронной сетью, он использует небольшой объем данных для тренировки модели и добивается высокой точности с наименьшим временем обучения.

Метод обучения свёрточной нейронной сети относится к обучению с учителем в машинном обучении, поэтому на этапе предварительной обработки изображения нам нужно создать обучающую выборку. Для этого нам необходимы исходные изображения (рисунок 2) и сегментированные.

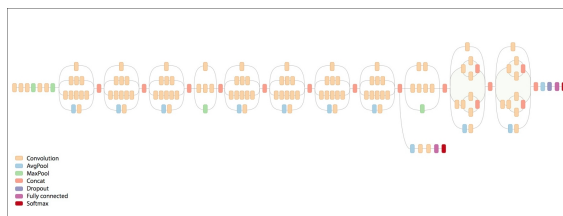


Рис. 1 – Архитектура модели Inception V3

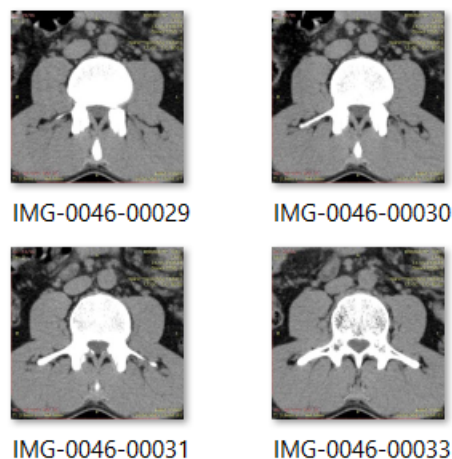


Рис. 2 – Пример исходных изображений

Для того, чтобы успешно переобучить нейронную сеть, необходимо изучить особенности строения позвонка человека и подготовить сегментированные изображения. Одновременно с этим требуется выбрать обозначения для каждой части позвонка.

Далее представлен перечень составных частей позвонка и их обозначение на итоговом изображении [3]:

- непарный остистый отросток (оранжевый);
- парные суставные отростки (желтый);
- ножка позвонка (зеленый);
- парные поперечные отростки (красный);
- тело позвонка (голубой).

Пример маркированных изображений представлен на рисунке 3.



Рис. 3 – Пример маркированных изображений, необходимых для обучения сети

В результате обучения сеть способна сегментировать изображения (рисунок 4).

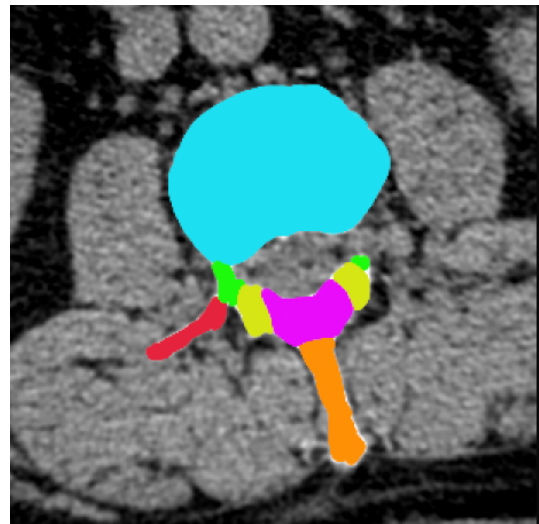


Рис. 4 – Пример сегментации, выполненной с помощью обученной модели

II. ЗАКЛЮЧЕНИЕ

В ходе данной работы была рассмотрена возможность сегментации медицинских изображений с помощью предварительно подготовленной модели Inception-v3. Разработанное приложение позволяет осуществлять поиск элементов позвонка на КТ-изображениях.

III. СПИСОК ЛИТЕРАТУРЫ

1. Bonnin, R. Building Machine Learning Projects with TensorFlow / R. Bonnin. – Birmingham: Packt Publishing, 2016. –267 p.
2. Fully convolutional networks for semantic segmentation / Long, J. [et al.]. // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) / IEEE. – Boston, 2015. – P. 3431–3440.
3. Билич, Г. Л. Атлас анатомии человека / Г. Л. Билич, В. Н. Николаенко. – Ростов н./Д.:Феникс – 488 с.

МОДЕЛЬ ОБРАБОТКИ СТРОК И СПИСКОВ ДАННЫХ ДЛЯ СИСТЕМ, УПРАВЛЯЕМЫХ ЗНАНИЯМИ

Ивашенко В. П.

Кафедра интеллектуальных информационных технологий, Факультет информационных технологий и управления, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ivashenko@bsuir.by

В статье рассматриваются команды и операции модели обработки строк и списков данных

ВВЕДЕНИЕ

При реализации систем, управляемых знаниями (СУЗ) [1], выделяются три уровня: уровень управления устройствами (ресурсами), уровень управления данными и уровень управления знаниями. Ранее были рассмотрены некоторые совместимые модели обработки информации [2] первых двух уровней, в том числе те, которые обеспечивают алгоритмическую универсальность [3]. В памяти с линейно организованным адресным пространством основными структурами данных (СД) с точки зрения реализации являются массивы, представляющие непрерывные структуры данных, и списки, которые могут храниться в памяти отдельными частями (множеством массивов). При работе с массивами, как и со списками, кроме основных, связанных с чтением и записью данных, требуется выполнять такие операции как перераспределение (реаллокация) и копирование данных. Если перечисленные операции для массивов в основном рассмотрены [4], то для списков некоторые их свойства по отношению к задаче управления данными остались невыявленными. Следует отметить, что одними из основных качеств интеллектуальных систем, к которым относятся СУЗ, являются обучаемость и открытость. При этом система должна обладать таким качеством как отзывчивость, т.е. системе следует быть или становиться доступной для ввода информации, в том числе для обучения, на протяжении непродолжительных временных интервалов, а не быть постоянно «самой себе на уме». С другой стороны, количество информации, которую нужно сообщить системе для того, чтобы её «обучить», не должно быть чрезмерным («учитель не должен решать задачу за ученика»), т.е. необходимо минимизировать количество сеансов ввода и объём «учебного материала». Эти требования можно сформулировать как требование минимизации при условиях, выраженных следующей системой.

$$\begin{cases} k * (1 + t) \geq k + n \\ t \geq 0 \\ k > 0 \\ n > 0 \\ \alpha > 0 \\ \beta > 0 \end{cases} \quad (1)$$

Пусть команда обрабатывается не более чем за время $1 + t$ ($t \geq 0$). Пусть, чтобы обработать данные за время n ($n > 0$), требуется k команд ($k > 0$), тогда оценка времени их работы не менее времени чтения команд и времени обработки данных: $k * (1 + t) \geq k + n$. Требуется минимизировать количество команд и максимально допустимое время их обработки.

$$\alpha * t + \beta * k \rightarrow \min \quad (2)$$

Решение этой задачи можно сформулировать как «принцип квадрата обучения»: квадрат количества команд (обучения) или время обработки каждой из них прямо пропорциональны времени, затраченному на решение задачи.

$$k = \sqrt{\frac{\alpha * n}{\beta}}; t = \sqrt{\frac{\beta * n}{\alpha}} \quad (3)$$

Исходя из этих условий и требования, отдельные операции, выполняемые системой, время их выполнения, по возможности должны соответствовать этому принципу.

Представление данных в виде списка может быть организовано для различных СД, в том числе логически доступных как строка или массив. Поэтому при работе с данными в виде списка можно выделить такие операции как доступ к элементу списка с указанным номером. Если элементами списка являются пары из элементов данных (ЭД), то это позволяет реализовать представление таких математических абстракций как функции и отношения на множестве ЭД соответствующего типа. Следует отметить, что далее рассматриваются списки, размер которых ограничен (разрядностью одной ячейки памяти): в случае недостатка свободной памяти операции не меняют исходные данные. Операции над списками неограниченного размера требуют отдельного рассмотрения.

ОПЕРАЦИИ ОБРАБОТКИ СТРОК И СПИСКОВ

Перечислим следующие операции над списками (строками): добавление элемента в начало списка, удаление элемента из начала списка, удаление всех элементов списка, конкатенация двух разных списков (строк), разделение списка (строки), доступ к элементу списка с указанным номером. Добавим к перечислению также

операции итерирования и операции над списками, хранящими для каждого ЭД его некоторый числовой индекс: операции добавления ЭД и его индекса, операции удаления ЭД и его индекса, поиска ЭД по его индексу. Кроме времени работы операции важной характеристикой её работы является её пропускная способность. Поэтому среди перечисленных операций нет операции копирования (создания) списка, но есть операция удаления элементов списка, так как она является критически важной с точки зрения пропускной способности: при удалении не создаётся никаких ЭД (кроме пустого списка). Для перечисленных операций за основу представления строк взята такая СД как «последовательностное дерево» (sequence tree) [5], которую можно рассматривать как модификацию такой СД как В-дерево. Время работы операций с этой структурой зависит от глубины дерева, которая в свою очередь зависит от количества хранимых элементов данных и количества ссылок в узле на другие узлы. Одной из самых затратных по времени операций является операция удаления элементов списка (строки), поэтому количество хранимых ссылок в узле определяется, исходя из соображений минимизации времени её работы. При выборе количества хранимых ссылок q ($q > 1$) близким к квадратному корню от (максимального) размера списка n , можно ограничить количество вызовов операции высвобождения памяти до $n/q \rightarrow \sqrt{n}$. Оценка времени и пропускной способности операций при этих условиях приведена в следующей таблице (см. табл. 1).

Таблица 1 – Характеристики операций

операция	временная сложность	оценка пропускной способности
push	$O(\log_q n * g(n))$	$O(\log_n q / g(n))$
pop	$O(h(r, q, n) + \log_q n * g(n))$	$O(1 / (h(r, q, n) + \log_q n * g(n)))$
popall	$O(h(r, q, n) + g(n) * n / q)$	$O(1 / (h(r, q, n) + g(n) * n / q))$
concat	$O(h(r, q, n) + \log_q n * g(n) + f(n) * n / q)$	$O(1 / (h(r, q, n) + \log_q n * g(n) + f(n) * n / q))$
valueat	$O(\log_2 n * f(n))$	$O(\log_n 2 / f(n))$
start increment decrement break	$O(h(r, q, n) + \log_q n * f(n))$	$O(1 / (h(r, q, n) + \log_q n * f(n)))$
over	$O(f(n))$	$O(1 / f(n))$
value	$O(f(n))$	$O(1 / f(n))$
keep	$O(\log_q n * g(n))$	$O(\log_n q / g(n))$
memorize lose	$O(h(r, q, n) + (\log_2^2 n + n / q) * f(n) + \log_q n * g(n))$	$O(1 / (h(r, q, n) + (\log_2^2 n + n / q) * f(n) + \log_q n * g(n)))$
remember	$O(h(r, q, n) + \log_2^2 n * f(n))$	$O(1 / (h(r, q, n) + \log_2^2 n * f(n)))$

Здесь g – функция временных затрат на перераспределение (высвобождение, выделение) памяти для структуры (строки) с n ЭД, f – функция временных затрат на доступ к од-

ной ячейке памяти, способной хранить структуру с n ЭД, h – функция временных затрат на обработку r итераторов над строкой длины n ($h(r, q, n) \sim r * f(r + n)$). Были проведены вычислительные эксперименты и сравнение результатов и характеристик работы (см. рис. 1) реализованных операций над представленными списками и «наивной» реализацией этих операций (<https://bitbucket.org/version/openjsvmm/>). Сравнение с аналогичными операциям над массивами JavaScript (IE 11) показало превосходство реализованных операций до четырёх раз.

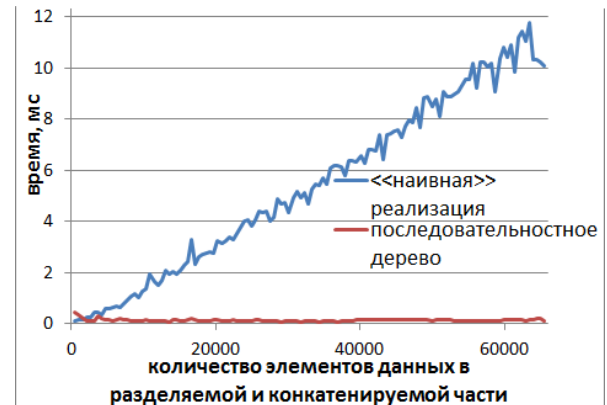


Рис. 1 – Сравнение времени выполнения операций разделения строки на две и их конкатенации

ЗАКЛЮЧЕНИЕ

Рассмотренные структуры, операции их спецификация и реализация, кроме непосредственного применения в задачах обработки данных, например при реализации пользовательского интерфейса, являются основой для реализации операций для обработки и представления знаний в соответствии с моделью унифицированного семантического представления знаний.

1. Голенков, В. В. Семантическая технология компонентного проектирования систем, управляемых знаниями / В. В. Голенков, Н. А. Гулякина // OSTIS. 2015. С. 57–78.
2. Ивашенко, В. П. Модели обработки информации в интеллектуальных системах, основанных на семантических технологиях / В. П. Ивашенко, А. С. Бельчиков, А. П. Еремеев // Информационные технологии и системы 2016 (ИТС 2016) : материалы международной научной конференции (БГУИР, Минск, Беларусь, 26 октября 2016) / редкол. : Л. Ю. Шилин [и др.]. — Минск: БГУИР, 2016. — С. 106–107.
3. Ивашенко, В. П. Модели обработки информации и программные средства для универсальных моделей решения задач / В. П. Ивашенко, // Информационные технологии и системы 2017 (ИТС 2017) : материалы международной научной конференции (БГУИР, Минск, Беларусь, 25 октября 2017) / редкол. : Л. Ю. Шилин [и др.]. — Минск: БГУИР, 2017. — С. 106–107.
4. Ивашенко, В. П. Операции управления массивами данных в линейно адресуемой памяти / В. П. Ивашенко, С. В. Синцов // Доклады БГУИР. – 2016. – № 6 (100). – С. 86–93.
5. Bates, R. Sequence-Trees: Slicing and Concatenation of Sequences in Logarithmic Time / R. Bates // 14th Midwest Conference on Combinatorics, Cryptography, and Comp. Wichita: WSU, – 2002. – Vol.42. –P. 33–60.

ИДЕНТИФИКАЦИЯ БИНАРНО ПОРОЖДАЕМЫХ СОБЫТИЙ ПРОЦЕССОВ ОБРАБОТКИ ЗНАНИЙ ДЛЯ СЕМАНТИЧЕСКОГО ПРОТОКОЛИРОВАНИЯ

Ивашенко В. П.

Кафедра интеллектуальных информационных технологий, Факультет информационных технологий и управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: ivashenko@bsuir.by

В статье рассматриваются принципы и алгоритм идентификации событий для семантического протоколирования на основе бинарно порождаемых событий

ВВЕДЕНИЕ

Семантическое протоколирование (СемПр) – механизм, призванный обеспечить интеллектуальную систему интроспективными возможностями с целью её надления качествами искусственного сознания, включая возможность объяснять собственную работу и её результаты [1]. СемПр заключается в записи на языке представления знаний, например в виде семантической сети, знаний о порядке (системе) выполняемых действий и событий (явлений) в процессах обработки знаний. С точки зрения обработки и управления данными СемПр тесно связано с таким свойством структур данных (СД) как персистентность. С точки зрения же СемПр спецификой является то, что рассматривается персистентность не СД отдельного типа, но СД, обладающей семантикой языка представления знаний, отражающей всю базу знаний интеллектуальной системы, с учётом распределённого характера работы с ней. Различают следующие виды персистентности СД [2]: частичную, когда доступны последовательно предшествующие версии СД и только последняя из них (текущая) доступна для модификации; полную, когда любую из последовательно предшествующих версий СД можно модифицировать (ветвящееся в будущее время); конфлюэнтную персистентность (персистентность со слиянием), когда пространственно-временная структура процессов обработки СД является ориентированным ациклически графом (DAG), и одна версия СД может иметь более одной предшествующей. При обработке знаний, однако, возможны процессы и с более сложной структурой, например, циклические, что может быть выявлено или обусловлено необходимостью интеграции знаний. Кроме того, работа с семантическим протоколом предполагает решение и других задач, кроме задач получения версии СД: задачи определения какое из событий (явлений) является предшествующим, какие явления (события) предшествуют каждому из заданных явлений, итерирование по предшествующим событиям и др. Ранее были предложены СД [1] и алгоритмы для решения задач СемПр [3], например, такая предложенная

структура как триплетно-свободный биномиальный стек [1] и соответствующие алгоритмы работы [3] с ней способна обеспечить полную персистентность структуры протоколируемых событий (явлений), а «пирамидальная структура» [1], структуры и алгоритм для протоколирования независимых событий [1] ориентированы на поддержку персистентности со слиянием.

ЗАДАЧИ СЕМАНТИЧЕСКОГО ПРОТОКОЛИРОВАНИЯ

В частном случае можно выделить процессы, в которых в (у)становлении каждого события участвуют не более пары событий. Такие события будем называть бинарно порождаемыми. В самом начале протокола находятся первичные события, которые не имеют устанавливающих событий. Каждое из последующих событий имеет двух предков, порождающих данное событие. Для каждого события может быть измерена длина – расстояние от самого дальнего первичного события по цепочке становлений. Событие и порождающие его события могут быть ассоциированы в грань в виде треугольника, в верхнем углу которого находится порождённое событие, а в углах при основании находятся порождающие события. Иерархическое продолжение построения ассоциаций в виде треугольных граней (получаемые грани аналогичным образом ассоциируются в грани более высокого уровня) позволяет получить пирамидальную структуру, в основании которой будет находиться структура, напоминающая структуру треугольника Серпинского. При конечном числе событий такая структура конечна. Грань верхнего уровня такой структуры будем называть вершиной (вершиной пирамидальной структуры), а количество уровней, на которых располагаются грани, будем называть высотой пирамидальной структуры. Для того чтобы различать события, с каждым событием, кроме вершины пирамидальной структуры, в которой находятся все события, предшествующие описываемому событию, ассоциируется некоторый идентификатор.

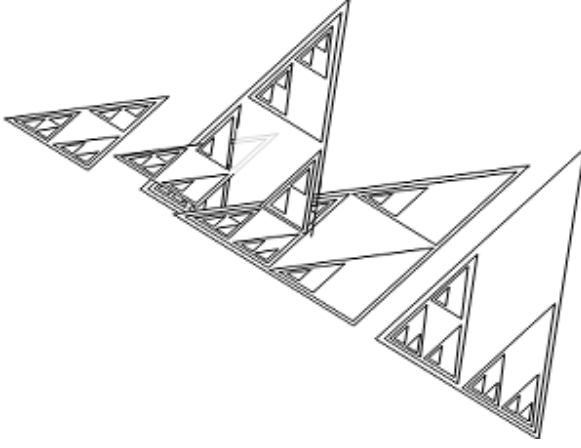


Рис. 1 – Схематическое представление семантического протокола

Особенностью подхода является то, что идентификатор каждого события (*EID*) формируется динамически по мере построения протокола в виде пирамидальных структур (см. Рис. 1), поэтому количество бит в идентификаторе (*Capacity*) пропорционально высоте сформированной пирамидальной структуры и называется «высотой идентификатора». Комбинированный идентификатор (*CID*) является результатом комбинации (объединения) идентификаторов (k, i) ассоциируемых гранью граней или событий (см. *combineCID* (1)). Так как каждый *EID* представляется цепочкой, то *CID* является структурой, соответствующей префиксному дереву идентификаторов событий, частью механизма индексации при поиске вхождения события в пирамидальную структуру.

```

combineCID
⟨c, k, i⟩ ← /* c – context (knowledge base) */
if (i ≠ ∅)
  if (k = ∅) k ← i
  else if (joined(⟨c, k, i⟩))
    |⟨c, k⟩ ← recombineCID(⟨c, k, i⟩)
  else
    |⟨c, t⟩ ← getCapacity(⟨c, i⟩)
    |⟨c, d⟩ ← getCapacity(⟨c, k⟩)
    |⟨c, f⟩ ← getFreeCapacity(⟨c, i⟩)
    |⟨c, b⟩ ← getFreeCapacity(⟨c, k⟩)
    |if (t ≤ b) ⟨c, k⟩ ← joinCID(⟨c, i, k, t, b⟩)
    else
      |if (d ≤ f) ⟨c, k⟩ ← joinCID(⟨c, k, i, d, f⟩)
      else
        |while (t – – > d)
        |  |⟨c, i⟩ ← setLower(⟨c, ∅, i⟩)
        |  |while (– – d > t)
        |  |  |⟨c, k⟩ ← setLower(⟨c, ∅, k⟩)
        |  |  |⟨c, k⟩ ← combineNode(⟨c, ∅, i, k⟩)
    |← ⟨c, k⟩

```

(1)

С каждой вершиной в дереве ассоциируется максимальная высота свободной ветви (*FreeCapacity*) в поддереве. При объединении деревьев, если высота одного из деревьев не превосходит максимальной высоты свободной вет-

ви в другом дереве, то это дерево помещается в качестве поддерева в узел ветви, которая становится занятой, максимальные высоты для узлов на пути от неё выше к корню соответственно обновляются. Иначе объединённое дерево строится путём добавления в качестве корня новой вершины, а в качестве поддеревьев в него добавляются объединяемые деревья (см. *joinCID* (2)).

```

joinCID
⟨c, i, k, t, b⟩ ←
⟨n, u⟩ ← ⟨k, k⟩
while (t ≤ b)
  |⟨f, p⟩ ← ⟨b, n⟩
  |⟨c, l⟩ ← getLower(⟨c, p⟩)
  |⟨c, n⟩ ← getUpper(⟨c, p⟩)
  |if (∅ ≠ l)
  |  |⟨c, b⟩ ← getFreeCapacity(⟨c, l⟩)
  |  |if (t ≤ b) n ← l
  |  |⟨c, b⟩ ← getFreeCapacity(⟨c, n⟩)
  |  |if ((b < f) ∧ (t ≤ b)) u ← p
  |while (t + + < b) ⟨c, i⟩ ← setLower(⟨c, ∅, i⟩)
  |⟨c, l⟩ ← getLower(⟨c, p⟩)
  |if (l = ∅) ⟨c, p⟩ ← setLower(⟨c, i, p⟩)
  |else ⟨c, p⟩ ← setUpper(⟨c, i, p⟩)
  |while (u ≠ p)
  |  |n ← u
  |  |⟨c, l⟩ ← getLower(⟨c, n⟩)
  |  |⟨c, u⟩ ← getUpper(⟨c, n⟩)
  |  |if (∅ ≠ l)
  |  |  |⟨c, b⟩ ← getFreeCapacity(⟨c, l⟩)
  |  |  |if (t ≤ b) u ← l
  |  |  |⟨c, u⟩ ← setFreeCapacity(⟨c, b, u⟩)
  |← ⟨c, k⟩

```

(2)

ЗАКЛЮЧЕНИЕ

Разработанные алгоритмы применимы для протоколирования процессов поддержки проектирования баз знаний, решения задач, включая обработку образов, в совместимых с технологией OSTIS [4] интеллектуальных системах.

1. Ивашенко, В. П. Алгоритмы семантического протоколирования процессов обработки знаний / В. П. Ивашенко // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3–4, 2018 / editorial board: M. Batura [etc.]. – Minsk, BSUIR, 2018. – P. 267–273.
2. Driscoll, J. R. Fully Persistent Lists with Catenation / J. R. Driscoll, D. D. Sleator, R. E. Tarjan. – 1994.
3. Ивашенко, В. П. Семантическое протоколирование процессов обработки знаний / В. П. Ивашенко // Информационные технологии и системы 2017 (ITS 2017) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) / редкол. : Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2017. – С. 110–111.
4. Голенков, В. В. Семантическая технология компонентного проектирования систем, управляемых знаниями / В. В. Голенков, Н. А. Гулякина // OSTIS. 2015. — С. 57—78.

ОНТОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ОТНОШЕНИЙ

Ивашенко В. П.

Кафедра интеллектуальных информационных технологий, Факультет информационных технологий и управления, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ivashenko@bsuir.by

В статье рассматриваются элементы онтологии пространственно-временных отношений, устанавливаемых на явлениях и событиях

ВВЕДЕНИЕ

В системах искусственного интеллекта одним из важных качеств является умение объяснить полученное решение, выразив его на языке, что является одним из признаков сознания [1]. Наличие соответствующего качества обеспечивает систему интроспективными возможностями с целью её совершенствования. Предлагается онтологическая модель пространственно-временных отношений, являющаяся основой для описания взаимосвязи явлений и событий в процессах обработки знаний [1]. На текущем этапе доступны известные темпоральные (временные) физические и абстрактные модели и соответствующие им логики, в рамках которых изучаются свойства темпоральных отношений [1-4]. Разработаны алгоритмы, позволяющие устанавливать различные отношения на моделях разной сложности (временные графы (timegraphs) и другие) [5-6]. Созданы системы, которые выявляют отношения в текстах и данных в глобальных компьютерных сетях [2,6]. Однако, разработанные модели темпоральных отношений ориентированы либо на описание явлений и процессов в конкретных физических системах в рамках проводимых экспериментов и недостаточно абстрактны для того, чтобы описать сложные информационные процессы, моделирующие поведение систем, обладающих сознанием, либо являясь абстрактными моделями, не приспособлены для описания информационных процессов, обладающих не только временными свойствами, но также и пространственной распределённостью в условиях присутствия НЕ-факторов знаний [1], включая их неполноту, неопределённость, неактуальность и гипотетичность [1]. В основу предлагаемого подхода положена онтологическая модель событий и явлений [1].

ОНТОЛОГИЯ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ОТНОШЕНИЙ

В онтологию пространственно-временных отношений входят понятия события и феномена (события или явления) и отношения между ними [1]. Приведём определяющие свойства основных из них на основе отношения (непосредственно) «прежде» (R_P или r).

Любое начинающее E событие феномена A ($BEG(E, A)$) есть событие, которое не имеет устанавливающих его событий феномена (т.е. находящихся прежде (R_P или r) него) ($\forall O ((\{O\} \subseteq A) \rightarrow (\neg R_P(O, E)))$). Любое оканчивающее E событие феномена A ($END(E, A)$) есть событие, которое не имеет устанавливаемых им событий феномена ($\forall O ((\{O\} \subseteq A) \rightarrow (\neg R_P(E, O)))$). Любое первичное (PRM) (финальное (FIN)) событие E феномена A есть событие, которое, если является оканчивающим (начальным) событием феномена без этого события, то принадлежит феномену, а если принадлежит феномену, то является начинающим (оканчивающим) событием феномена без этого события (1).

$$\begin{aligned} & ((END(E, (A/\{E\})) \rightarrow (\{E\} \subseteq A)) \\ & \wedge ((\{E\} \subseteq A) \rightarrow BEG(E, (A/\{E\}))))); \quad (1) \\ & ((BEG(E, (A/\{E\})) \rightarrow (\{E\} \subseteq A)) \\ & \wedge ((\{E\} \subseteq A) \rightarrow END(E, (A/\{E\})))) \end{aligned}$$

Любое событие E переустановки (PIB) события установки феномена A есть событие, которое устанавливает все события, кроме первичных событий феномена, установленные событием установки феномена (2).

$$\begin{aligned} & (PRM(O, A) \wedge (\forall T ((\{T\} \subseteq A) \rightarrow (PRM(T, A) \\ & \vee (R_P(E, T) \rightarrow R_P(O, T)))))) \quad (2) \end{aligned}$$

Любое событие E предустранения (FIF) события устранения феномена A есть событие, которое устанавливается всеми событиями, кроме финальных событий феномена, устанавливающими событие устранения феномена (3).

$$\begin{aligned} & (FIN(O, A) \wedge (\forall T ((\{T\} \subseteq A) \rightarrow (FIN(T, A) \\ & \vee (R_P(T, E) \rightarrow R_P(T, O)))))) \quad (3) \end{aligned}$$

Любое событие E появления (OCR) (уничтожения (DES)) феномена A есть событие, которое является первичным (финальным) событием феномена и для любого события, являющегося первичным (финальным) событием и событием переустановки (предустранения) данного события, если оно находится позже (раньше) данного, то,

если и только если оно и есть данное (4).

$$\begin{aligned} & (PRM(E, A) \wedge (\forall O(PIB(A, E, O) \\ & \rightarrow ((R_B(E, O) \rightarrow (\{E\} = \{O\})))))); \\ & (FIN(E, A) \wedge (\forall O(FIF(A, E, O) \\ & \rightarrow ((R_B(O, E) \rightarrow (\{E\} = \{O\})))))) \end{aligned} \quad (4)$$

Любое событие E установления (INS) (удаления (RMV)) феномена A есть событие, которое является первичным (финальным) событием феномена и не является событием появления (уничтожения) феномена, если и только если любое первичное (финальное) событие, являющееся событием переустановки (предустранения) данного события, не является событием появления (уничтожения) феномена (5).

$$\begin{aligned} & (PRM(E, A) \wedge (\forall E(\neg OCR(E, A) \\ & \sim (\forall O(PIB(A, E, O) \rightarrow (\neg OCR(O, A)))))); \\ & (FIN(E, A) \wedge (\forall E(\neg DES(E, A) \\ & \sim (\forall O(FIF(A, E, O) \rightarrow (\neg DES(O, A)))))) \end{aligned} \quad (5)$$

Любое первое (FST) (последнее (LST)) событие E феномена A есть событие феномена, являющееся начинающим (оканчивающим) событием феномена $((\{E\} \subseteq A) \wedge BEG(E, A))$ $((\{E\} \subseteq A) \wedge END(E, A))$.

Любое примыкающее (IFW) (примкнутое (IBW)) явление E к A есть явление, любое событие после (прежде) любого события которого принадлежит другому явлению (6).

$$\begin{aligned} & (\forall T(\{T\} \subseteq E) \\ & \rightarrow (\forall O(R_P(T, O) \rightarrow (\{O\} \subseteq A))); \\ & (\forall T(\{T\} \subseteq E) \\ & \rightarrow (\forall O(R_P(O, T) \rightarrow (\{O\} \subseteq A)))) \end{aligned} \quad (6)$$

Любое событие E или явление непосредственно перед (PRV) событием E или явлением, если и только если первое прежде второго или первое примыкает к объединению первого и второго, второе примкнуто к объединению первого и второго, а первое вне второго (7).

$$\begin{aligned} & (R_P(E, A) \vee ((IFW(E, (E \cup A)) \\ & \wedge IBW(A, (E \cup A))) \wedge R_O(E, A)) \end{aligned} \quad (7)$$

Любое событие E или явление ранее (BFR) события или явления A , если и только если первое прежде второго или первое раньше второго или неверно, что любое событие, которое после первого, не ранее второго (8).

$$\begin{aligned} & ((R_P(E, A) \vee R_B(E, A)) \vee (\neg(\forall O(R_P(E, O) \\ & \rightarrow (\neg BFR(O, A))))) \end{aligned} \quad (8)$$

Основными отношениям, которые рассматривает онтологическая модель являются: (непосредственно) «прежде» (R_P), (непосредственно) «после» (R_N), «раньше» (R_B), «позже» (R_A), «с» (R_S), «по» (R_T), «от» (R_F), «до» (R_U), «когда» (R_H), «тогда» (R_E), «в течение» (R_W), «во время» (R_D), «на протяжении» (R_V), «в» (R_I), «вне» (R_O), R_X – «не определено».

Таблица 1 – Пространственно-временные отношения

предикат	определяющее выражение
$R_P(A, O)$	$PRV(A, O)$
$R_N(A, O)$	$R_P(O, A)$
$R_B(A, O)$	$(BFR(A, O) \vee (\forall T(\{T\} \subseteq E) \rightarrow (\forall O(\{O\} \subseteq A) \rightarrow BFR(T, O))))$
$R_A(A, O)$	$R_B(O, A)$
$R_{BQ}(A, O)$	$(R_B(A, O) \vee (\{A\} = \{O\}))$
$R_{AQ}(A, O)$	$R_{BQ}(O, A)$
$R_S(A, O)$	$((\forall E(INS(E, A) \sim INS(E, O))) \vee (\{A\} = \{O\}))$
$R_T(A, O)$	$((\forall E(RMV(E, A) \sim RMV(E, O))) \vee (\{A\} = \{O\}))$
$R_F(A, O)$	$((\forall E(\{E\} \subseteq A) \rightarrow (\neg(\forall T(\{T\} \subseteq O) \rightarrow (\neg R_{BQ}(T, E))))) \vee R_{BQ}(O, A))$
$R_U(A, O)$	$((\forall E(\{E\} \subseteq A) \rightarrow (\neg(\forall T(\{T\} \subseteq O) \rightarrow (\neg R_{AQ}(T, E))))) \vee R_{AQ}(O, A))$
$R_W(A, O)$	$(R_F(A, O) \wedge R_U(A, O))$
$R_H(A, O)$	$R_W(A, O)$
$R_E(A, O)$	$R_H(O, A)$
$R_D(A, O)$	$((\neg R_U(O, A)) \wedge (\neg R_F(O, A)))$
$R_V(A, O)$	$((\neg R_B(O, A)) \wedge (\neg R_A(O, A)))$
$R_I(A, O)$	$((\forall E(\{E\} \subseteq A) \rightarrow (\{E\} \subseteq O)) \vee (\{A\} = \{O\}))$
$R_O(A, O)$	$((\forall E(\{E\} \subseteq A) \rightarrow (\neg(\{E\} \subseteq O))) \wedge (\neg(\{A\} = \{O\})))$

ЗАКЛЮЧЕНИЕ

Понятия, введенные в предложенной онтологической модели, предназначены и применяются для описания процессов обработки знаний в базах знаний, использующих средства модели унифицированного семантического представления знаний в сочетании с такими моделями, механизмами и методами их обработки как продукции, процедуры и нейронные сети с целью семантического протоколирования для объяснения результатов процессов обработки знаний [1].

1. Ивашенко, В. П. Алгоритмы семантического протоколирования процессов обработки знаний / В. П. Ивашенко // BIG DATA Advanced Analytics: collection of materials of the fourth international scientific and practical conference, Minsk, Belarus, May 3-4, 2018 / editorial board: M. Batura [etc.]. – Minsk, BSUIR, 2018. – P. 267–273.
2. Derczynski, L., Gaizauskas, R. Empirical Validation of Reichenbach's Tense Framework. Proceedings of the 10th Conference on Computational Semantics, ACL. – 2013. – pp. 71–82.
3. Staab, S. From binary temporal relations to non-binary ones and back. Artificial Intelligence. – 2001. – Vol. 128 (1-2). – pp. 1–29.
4. Pustejovsky, J., Ingria, B., Sauri, R., Castano, J., Littman, J., Gaizauskas, R., Setzer, A., Katz, G., Mani, I. The specification language TimeML. In: The Language of Time: A Reader, Oxford University Press. – 2005. – pp. 545–557.
5. Gerevini, A., Schubert, L. K. Efficient Temporal Reasoning through Timegraphs. In: Proceedings of the 13th International Joint Conference on Artificial Intelligence, California, USA. – 1993. – Vol. 1. – pp. 648–654.
6. Llorens, H., Chambers, N., UzZaman, N., Mostafazadeh, N., Allen, J. F., Pustejovsky, J.: SemEval-2015 Task 5: QA TempEval – Evaluating Temporal Information Understanding with Question Answering. – 2015. – pp. 792–800.

НОВЫЕ ЗАДАЧИ УПРАВЛЕНИЯ ИСПОЛЬЗОВАНИЕМ РАДИОЧАСТОТНОГО СПЕКТРА И ПУТИ ИХ РЕШЕНИЯ

Карпук А. А.

Кафедра программного обеспечения сетей телекоммуникаций,

Белорусская государственная академия связи

Минск, Республика Беларусь

E-mail: A_Karpuk@mail.ru

Обоснована необходимость решения новых задач в информационных системах управления использованием радиочастотного спектра. Показаны возможные подходы к решению новых задач

ВВЕДЕНИЕ

Практически все информационно – коммуникационные технологии и системы в качестве коммуникационной среды используют радиочастотный спектр (РЧС), под которым понимается совокупность радиочастот, которые могут быть использованы для функционирования радиосредств или высокочастотных устройств. РЧС является ограниченным однократно используемым возобновляемым природным пространственно – временным ресурсом. Указанные обстоятельства приводят к необходимости управления использованием РЧС, которое производится на международном, региональном и национальном уровнях. Наиболее сложные задачи управления использованием РЧС решаются на национальном уровне национальными Администрациями радиосвязи [1]. Для автоматизированного решения этих задач создаются и используются информационные системы управления (ИСУ) использованием РЧС [2]. Одной из важнейших задач ИСУ использованием РЧС является задача анализа возможности присвоения номиналов или полос радиочастот по заявкам пользователей.

I. НОВЫЕ ЗАДАЧИ УПРАВЛЕНИЯ ИСПОЛЬЗОВАНИЕМ РЧС

В существующих ИСУ использованием РЧС анализ возможности присвоения номиналов или полос радиочастот по заявкам пользователей производится путем проверки соответствия заявленных частотных присвоений национальным планам использования радиочастот и проверки выполнения норм частотно - территориального разнеса, указанных в нормативных документах или в технических характеристиках РЭС. Нормы частотно - территориального разнеса проверяются по совпадающим и соседним радиоканалам, остальные побочные и интермодуляционные излучения передатчиков и побочные и интермодуляционные каналы приема приемников не учитываются. Методы проверки выполнения норм частотно - территориального разнеса не учитывают реальной рефракции радиоволн в атмосфере, дифракционных потерь на рельефе местности, потерь распространения в за-

стройке и растительности, потерь распространения из-за метеорологических факторов. Это приводит к весьма грубой оценке возможности присвоения номиналов или полос радиочастот по заявкам в сторону ее занижения, что значительно снижает эффективность использования РЧС.

Для устранения перечисленных недостатков в ИСУ использованием РЧС должны решаться новые задачи. Требуется изменить подход к решению задач анализа возможности присвоения номиналов или полос радиочастот по заявкам пользователей РЧС. Вместо проверки выполнения норм частотно - территориального разнеса нового частотного присвоения с имеющимися частотными присвоениями предлагается решать новую задачу оценки качества радиосвязи на заявленных номиналах или полосах радиочастот и оценки влияния работы РЭС на заявленных номиналах или полосах радиочастот на работу других РЭС с ранее присвоенными радиочастотами. При решении этой задачи должны учитываться пространственное положение РЭС, технические характеристики РЭС и антенн, реальная рефракция радиоволн в атмосфере, дифракционные потери на рельефе местности, потери распространения в застройке и растительности, потери распространения из-за метеорологических факторов. Должны учитываться основное, внеполосное, побочные и интермодуляционные излучения передатчиков и основной, внеполосные, побочные и интермодуляционные каналы приема приемников.

В результате решения задачи определяются возможные электромагнитные помехи, и для каждой помехи вычисляется уровень (величина) помехи, который не будет превышен в течение заданного процента времени (например, в течение 99,9 процента времени). На основе анализа уровней помех принимается решение об их допустимости и о возможности выдачи пользователям разрешений на использование радиочастот.

Национальные Администрации радиосвязи должны оказывать пользователям услугу по выбору номиналов или полос радиочастот, для которых возможно получение разрешения на использование. Для оказания такой услуги национальная Администрация радиосвязи должна ре-

шить новую задачу выбора заданного количества номиналов или полос радиочастот для работы РЭС с заданными характеристиками и заданным пространственным положением. Эта задача решается для новых РЭС в предположении, что для РЭС с ранее присвоенными радиочастотами известны их технические характеристики и пространственное положение.

В районах крупных мегаполисов ресурсы РЧС в наиболее используемых частотных диапазонах практически полностью исчерпаны. Дальнейшее развитие систем радиосвязи возможно только за счет освоения новых частотных диапазонов или оптимизации использования номиналов и полос радиочастот в существующих частотных диапазонах. При оптимизации использования радиочастот производится изменение присвоенных номиналов и полос радиочастот для РЭС с ранее присвоенными радиочастотами. К примеру, в Российской Федерации планируется только для подвижных радиослужб к 2020 г. увеличить эффективность использования выделенных полос РЧС на 35% за счет решения новой задачи оптимизации присвоения радиочастот.

Заметим, что задача выбора заданного количества номиналов или полос радиочастот для работы РЭС с заданными характеристиками и заданным пространственным положением является частным случаем задачи оптимизации присвоения номиналов и полос радиочастот для РЭС, решаемой в предположении, что номиналы и полосы радиочастот ряда РЭС фиксированы и не подлежат изменению. Поэтому в дальнейшем можно рассматривать две новые задачи: задачу оценки качества радиосвязи на заявленных номиналах или полосах радиочастот и задачу оптимизации присвоения номиналов и полос радиочастот для работы РЭС.

II. ПУТИ РЕШЕНИЯ НОВЫХ ЗАДАЧ

Для решения новых задач управления использованием РЧС требуются глубокие знания о предметной области сетей радиосвязи. В базе знаний должны содержаться знания о технических характеристиках и координатах РЭС, о технических характеристиках и составе радиолиний и радиосетей, о топографических характеристиках местности (рельефе местности, координатах и типах застройки, координатах лесных массивов и водных поверхностей, электромагнитных свойствах почвы), о метеорологических характеристиках местности (рефракционных свойствах атмосферы, вероятности и интенсивности осадков, тумана, облаков, атмосферных газов и водяных паров) в районе размещения РЭС. При решении задач должны использоваться знания о трассах распространения радиоволн и влиянии технических характеристик РЭС, топографических и метеорологических характеристик местности на распространение радиоволн.

В настоящее время для построения формализованного описания предметной области информационных систем используются методология IDEF1X и модель «сущность – связь» П. Чена. В предметной области ИСУ использованием РЧС многие сущности имеют иерархическую структуру, в качестве значений атрибутов могут выступать элементы данных, векторы, структуры и повторяющиеся группы элементов, векторов и структур данных. Это делает невозможным использование традиционных методологий. Кроме того, традиционные методологии построения описания предметной области направлены на проектирование структур данных, а не на выявление связей между данными и представление знаний о предметной области.

Для получения формализованного описания предметной области ИСУ использованием РЧС предлагается построить онтологию предметной области. В качестве множества концептов этой онтологии следует использовать понятия из таблиц данных по управлению использованием спектра, приведенных в справочнике [2]. Путем выделения и описания отношений между концептами типа система – элемент, род – вид, объект – свойство и т. д. можно построить иерархическую структуру сущностей предметной области. Функции интерпретации, заданные на сущностях и отношениях онтологии, реализуются в виде модулей и библиотек программ, которые могут разрабатываться разными специалистами и добавляться в онтологию. В начальную версию онтологии можно включить модули и библиотеки программ системы оценки качества радиосвязи и оптимизации присвоения частот радиолиниям, описанные в монографии [3].

Задача оценки качества радиосвязи на заявленных номиналах или полосах радиочастот может решаться на основе онтологии предметной области с использованием модулей и библиотек программ, реализующих функции интерпретации. Задача оптимизации присвоения номиналов и полос радиочастот для работы РЭС также должна решаться на основе онтологии предметной области с использованием математической модели задачи [3] и метаэвристических приближенных алгоритмов.

III. СПИСОК ЛИТЕРАТУРЫ

1. Справочник по управлению использованием спектра на национальном уровне. Издание 2015 года. – МСЭ-Р, 2015. – 364 с.
2. Справочник по компьютерным технологиям управления использованием радиочастотного спектра (КТ). Издание 2015 года. – МСЭ-Р, 2015. – 192 с.
3. Карпук, А.А. Системы оценки качества радиосвязи и оптимизации присвоения частот радиолиниям: монография / А.А. Карпук. – Воронеж: Научная книга, 2015. – 230 с.

ВЫДЕЛЕНИЕ ОСНОВНОЙ ДОРОЖНОЙ СЕТИ НА ИЗОБРАЖЕНИЯХ ЗЕМНОЙ ПОВЕРХНОСТИ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ U-NET

Кочурко А. А., Абламейко С. В.

Кафедра веб-технологий и компьютерного моделирования, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: aliaksandr.kachurka@gmail.com, ablameyko@bsu.by

В работе описано решение задачи автоматического построения карты дорог по снимкам земной поверхности высокого разрешения на основе искусственной нейронной сети, а также предложен алгоритм построения обучающего множества для данной задачи. Использование нейронной сети обусловлено простотой формирования обучающего множества: большие объёмы изображений земной поверхности и соответствующих дорожных карт находятся в открытом доступе.

ВВЕДЕНИЕ

Карта дорожной сети – ключевой компонент в работе многих важных сервисов. Служба скорой помощи нуждается в точной дорожной карте для планирования пути к месту вызова; совокупность системы глобального позиционирования и карты дорог делает возможным существование программ, предоставляющих пользователю информацию об оптимальном маршруте к месту назначения.

Поскольку дорожная сеть постоянно изменяется, возникает проблема поддержания дорожных карт в актуальном состоянии. В настоящее время карты дорожной сети строятся и обновляются по снимкам земной поверхности высокого разрешения вручную [1]. Процесс обновления карт, очевидно, затратный по времени. Поэтому есть необходимость в системе, способной автоматически строить дорожные карты.

I. ОПИСАНИЕ РЕШЕНИЯ

При решении задач обработки изображений часто используются *свёрточные нейронные сети (СНС)* [2]. СНС хорошо справляются с задачей классификации объектов на изображениях, когда входному изображению ставится в соответствие метка класса, к которому, вероятнее всего, принадлежит объект [1]. Однако во многих задачах, связанных с обработкой изображений, желаемый выход включает также информацию о локализации, т.е. метка класса должна присваиваться каждому пикселю входа. Поэтому в работе для решения задачи использовался экземпляр модифицированной свёрточной нейронной сети, имеющий структуру *U-Net*.

Идея, заложенная в структуру *U-Net*, заключается в замене полносвязных слоёв СНС на дополнительные пары свёрточный слой-слой повышающей дискретизации (ПД). Добавленные слои называются расширяющими, оставшиеся слои исходной СНС – сжимающими. Выход слоя ПД комбинируется с картами признаков более высокого разрешения сжимающих слоёв. Таким

образом, свёрточный слой, следующий за слоем повышающей дискретизации, получает информацию, позволяющую делать вывод не только о том, *что* изображено, но и о том, *где* [1].

Используемый экземпляр нейронной сети состоит из 18 слоёв.

Размерность входа нейронной сети – 160×160 .

Количество связей с изменяемыми весами – 7851537.

Формирование обучающего множества оказалось ключевым этапом в решении рассматриваемой задачи. Алгоритм создания ОМ включает 2 шага:

1. преобразование исходных ВКТ-данных к растровой маске дорог [3];
2. выбор фрагментов изображений и соответствующих масок.

Шаг 2 всегда имеет место при формировании обучающего множества для *U-Net* и обычно заключается в выборе некоторого числа произвольных фрагментов каждого из исходных изображений. Однако, исходя из особенностей конкретной задачи, можно сформулировать алгоритм выбора фрагментов. Составление такого алгоритма представляет интерес.

Дороги покрывают лишь малую часть земной поверхности. При неудачном выборе фрагментов может быть сформировано обучающее множество, в котором дороги практически не будут представлены. В рамках работы сформулирован алгоритм, позволяющий избежать такой ситуации:

1. Исходное изображение размера $X \times Y$ и соответствующая маска M делятся на четверти.
2. Для каждой четверти C по соответствующей части маски M_C подсчитывается *коэффициент заполнения*

$$k = \frac{4 \sum_{i=0}^{\frac{X}{2}} \sum_{j=0}^{\frac{Y}{2}} M_C(i, j)}{XY}$$

Четверть с большей площадью дорог будет иметь больший коэффициент k .

3. Количество l_C сегментов, которые следует взять в четверти C , вычисляется по формуле

$$l_c = Pk$$

где значение $P = 3000$ выбрано опытным путём исходя из нужного числа сегментов.

Вместе с выбранными фрагментами в обучающее множество были включены их копии, отражённые по вертикали и горизонтали. Таким образом, объём множества увеличился в 4 раза.

В результате было сформировано обучающее множество с 12256 элементами, каждый из которых представляет собой пару фрагмент исходного изображения - соответствующий сегмент маски. Размер фрагментов 160×160 .

Обучение проводилось методом обратного распространения ошибки в 30 итераций по обучающему множеству.

II. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Реализованная программа способна выделять дорожную сеть на 3-канальных изображениях земной поверхности. Распознавание дорог осуществляется с помощью обученного экземпляра нейронной сети U-Net, описанного ранее.

Входное изображение делится на фрагменты размера, совпадающего с размером входа нейронной сети (160×160). Каждый фрагмент обрабатывается экземпляром U-Net, после чего полученные в результате обработки маски объединяются в общую маску дорог для исходного изображения.

Программа написана на языке *Python 3.6.3* с использованием библиотеки *Keras* на основе *Theano* [5].

III. РЕЗУЛЬТАТЫ

Одной из метрик, характеризующих успешность работы нейронной сети, является *точность* [4]. Обучающее множество, подготовленное с использованием описанного алгоритма, позволило добиться большей точности при обучении, чем при произвольном выборе фрагментов изображений.

Таблица 1 – Сравнение результатов обучения различных конфигураций нейронной сети на различных обучающих множествах

Конфигурация НС	Кол-во каналов входных данных	Алгоритм формирования ОМ	Точность
[3]	16	Произвольный	0.94
Описанная конфигурация	3	Произвольный	0.84
Описанная конфигурация	3	Описанный алгоритм	0.91

Полученные результаты для 3-канальных изображений оказались лучше, чем результаты

решения [3] для 3-канальных снимков, но уступают результатам [3] на 16-канальных снимках. Однако следует заметить, что 16-канальные изображения содержат гораздо больше информации, а значит и обучение на таких снимках более эффективно.

Также видно, что при использовании разработанного алгоритма выбора фрагментов изображения на одних и тех же исходных данных удалось сгенерировать более качественное обучающее множество, т.е. множество, на котором нейронная сеть за одинаковое число итераций обучилась лучше, чем на других.

Далее приведены наглядные результаты работы реализованной программы на примере снимка земной поверхности высокого разрешения.

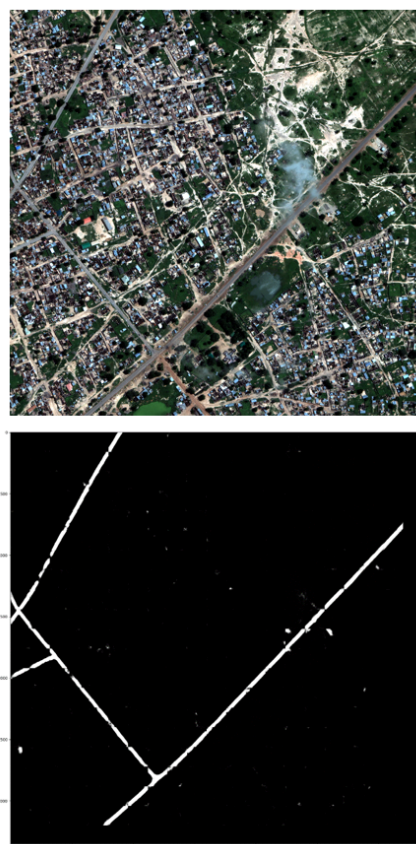


Рис. 1 – Пример входных данных (выше) и результат работы программы

1. Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox. – 8 p.
2. Goodfellow, I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – 785 p.
3. Dstl-Satellite-Imagery-Detection [Electronic resource]. Mode of access: <https://github.com/DeepVoltaire/Dstl-Satellite-Imagery-Feature-Detection>. – Date of access: 11.03.2018.
4. Kriesel, D. A Brief Introduction to Neural Networks / D. Kriesel. – 226 p.
5. Theano documentation [Electronic resource]. Mode of access: <http://deeplearning.net/software/theano>. – Date of access: 11.11.2017.

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ УПРАВЛЕНИЯ РОБОТОМ ИЗ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Курочка К. С., Оныськив В. В.

Кафедра «Информационные технологии», Гомельский государственный технический университет им. П.О. Сухого

Гомель, Республика Беларусь

E-mail: kurochka@gstu.by, oniskif.vladislaw@gmail.com

Рассмотрено проектирование аппаратного и программного обеспечения колесного робота и программного обеспечения для управления такими роботами из виртуальной реальности используя мобильные устройства.

ВВЕДЕНИЕ

На сегодняшний день информационные компьютерные системы все больше прогрессируют и развиваются с каждым днем, а вместе с ними идет развитие связанной с ней другой системы – виртуальной реальности. Виртуальная реальность – это интерактивная имитация реалистичных и вымышленных сред, т.е. некий иллюзорный мир, в который погружается и с которым взаимодействует человек [1].

Не стоит на месте, а также развивается быстрыми темпами и такая наука как робототехника (наука об архитектуре, создании и управлении роботами). В связи с растущей актуальностью использования роботов в современном мире особую важность приобретает разработка программного обеспечения в этом направлении.

I. ОПИСАНИЕ СХЕМЫ ПРОЕКТИРОВАНИЯ РОБОТА

В рамках данной работы ставится задача разработать эффективное программное средство по управлению роботом из виртуального пространства. Проект включает в себя три основных независимых компонента (рис. 1): колесный робот, сервер и мобильный телефон, проектирование и разработка программного обеспечения для которых осуществляется отдельно.

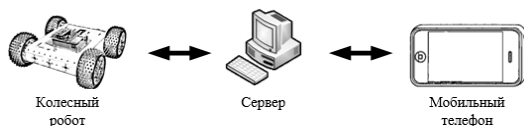


Рис. 1 – Основные компоненты проекта

Колесный робот состоит из нескольких основных блоков (рис. 2):

1. Блок связи и «мозг робота» – за основу взят одноплатный компьютер Orange Pi PC Plus. Данный элемент представляет из себя полноценный компьютер миниатюрных размеров, позволяющий работать со множеством UNIX – подобных операционных систем. Блок связи выполняет функцию обработчика, он получает команды управления, декодирует/распаковывает их и пере-

дает на блок управления, а также кодирует/запаковывает данные от видеокamеры и отправляет их на сервер в сеть Интернет.

2. Видеокamera – веб-камера или модуль камеры совместимый с Orange Pi, используемый для получения и дальнейшей передачи видео потока пользователю.
3. Блок управления – драйвер двигателя L298N. Данный драйвер построен на базе микросхемы работающей по принципу H-моста. В следствии чего, мы имеем возможность независимого управления двумя двигателями и в разных направлениях [2].
4. Двигатели – два мотора с рабочим напряжением 6-12В.
5. Блок питания – повербанк выдающий напряжение 5В с силой тока 2.5А.

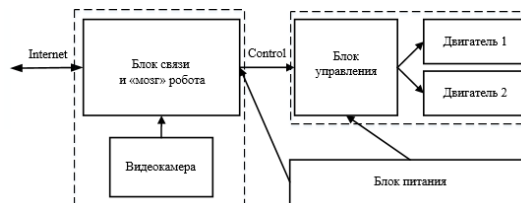


Рис. 2 – Основные компоненты проекта

Все перечисленные элементы в совокупности представляют полноценного колесного робота, который принимает команды извне, декодирует их и отправляет на блок управления, в следствии чего приводятся в движение двигатели. Параллельно работает видеокamera, она захватывает видеопоток, отправляет его на «мозг» робота для обработки и кодирования с дальнейшей передачей на сервер в сети Интернет. Создание программного обеспечения модуля управления и связи производится на языках C/Python и представляет из себя команды для управления подаваемые на порты GPIO и команды работы с сетью для приема команд управления и передачи видео [3].

В качестве мобильного телефона может использоваться любой мобильный телефон на базе OS Android версии 4.4 и выше. Разработка программного обеспечения происходит на языке Java с использованием пакета разработки Google

VR SDK. Пользователь, управляя роботом с помощью такого программного обеспечения, выполняет различные действия: повороты головы, наклоны и т.д. Эти действия трансформируются в команды управления с дальнейшей отправкой на сервер.

Связующим звеном между роботом и мобильным устройством выступает веб-сервер, расположенный на хостинге в сети Интернет. Данное программное обеспечение не выполняет никаких действий по обработке данных, а лишь переадресует данные от одного элемента к другому. Поскольку сервер находится в сети Интернет, мы не имеем ограничений в несколько метров, которые устанавливает работа по локальной сети с использованием технологии Wi-Fi, а имеем безграничные возможности по расстоянию и управлению роботом за тысячи километров.

II. ОБЛАСТИ ПРИМЕНЕНИЯ

Разработка новых интерфейсов для управления роботом в виртуальной среде является весьма перспективным и, в тоже время, весьма необходимым направлением. С первого взгляда, не сразу удастся представить сферы использования такой технологии, однако поразмыслив, они очень широки. В первую очередь, это труднодоступные или совсем недоступные для человека места в связи с физическими или атмосферными факторами. К сферам использования таких роботов можно отнести космические и глубоководные исследования, ликвидации последствий техногенный катастроф и аварий, использование в сфере вооруженных сил, выполнение сложнейших медицинских операций, использование в беспилотных малогабаритных летательных аппаратах.

III. ПОТЕНЦИАЛЬНЫЕ ПРОБЛЕМЫ

Однако, разработка таких интерфейсов не обделена ограничениями и трудностями. Самым главным ограничением и условием для комфортной работы является наличие хорошего высокоскоростного подключения к сети Интернет. Основной проблемой управления через Интернет является наличие временных задержек при передаче данных, и чем больше расстояние, тем эти задержки больше. Это делает управление через Интернет затруднительным, а во многих случаях невозможным. Так как видеопоток, получаемый с камеры робота, имеет достаточно большой объем данных и, следовательно, необходим очень хороший канал связи через Интернет. Для преодоления этой проблемы, сейчас остается весьма актуальной создание инновационных методов по повышению эффективности управления роботами через Интернет.

Одним из способов решить эту проблему является уменьшение качества передаваемого изображения. Однако, из этого вытекает следующая проблема, а именно появление неудобств управления для оператора. Так как качество изображения становится хуже, то и распознавать объекты, оценивать положение робота и расстояние между объектами становится гораздо сложнее, что может привести к совершенно непредсказуемым последствиям.

Другим вариантом преодоления проблемы является использование смоделированных виртуальных трехмерных моделей вместо реальных объектов, которые окружают рабочее пространство робота. Идея построена на том, что вместо большого по объему видеопотока можно передавать минимальный набор параметров, однозначно определяющих состояние робота и его рабочей среды (набор координат местоположения робота, координаты наблюдаемого объекта для взаимодействия).

ЗАКЛЮЧЕНИЕ

Но технологии не стоят на месте. Со временем крупные компании и ученые улучшат каналы связей, будут придуманы новые методы передачи данных, улучшатся и сами устройства виртуальной реальности и доработаются интерфейсы, в силу чего VR-инструменты позволят нам взаимодействовать с объектами так, как никогда раньше.

Конечно, нам потребуется около пары поколений, чтобы внедрить технологии виртуальной реальности во многие сферы жизни человека, однако польза от таких технологий будет не меньше чем все трудности, которые предстоит пройти. Виртуальная реальность – это новая развивающаяся среда, которая способна оказать на общество огромный эффект, сравнимый с появлением радио или телевидения. И сегодня мы в силах заложить фундаменты в разработке интерфейсов и опыта взаимодействия для этой среды без того, чтобы по умолчанию применять уже существующие решения.

СПИСОК ЛИТЕРАТУРЫ

1. Виртуальная реальность [Электронный ресурс] / – Режим доступа: <https://ru.wikipedia.org/wiki/>. – Дата доступа: 10.08.2018.
2. Управляем шаговыми двигателями и DC моторами, L298. [Электронный ресурс] / – Режим доступа: <https://ph0en1x.net/87-l298-and-raspberry-pi-stepper-and-dc-motor.html>. – Дата доступа: 12.08.2018.
3. Робот на RaspberryPi, Arduino и OpenCV. [Электронный ресурс] / – Режим доступа: <https://habr.com/post/249421>. – Дата доступа: 10.08.2018.

СМАРТФОН КАК ИНСТРУМЕНТ ДЛЯ ПРЕДОСТАВЛЕНИЯ ПЕРСОНАЛЬНЫХ РЕКОМЕНДАЦИЙ СТРАТЕГИИ СБАЛАНСИРОВАННОГО ПИТАНИЯ

Лапицкая Н. В., Трус В. В., Ильюкевич А. В., Антоненко Д. А., Варфоломеев А. В
Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: lapan@bsuir.by

Целью этой работы является разработка инструмента для сбора информации о дневнике питания пользователя и основы для формирования базы знаний для разработки системы рекомендаций по личной стратегии здорового питания на основе предпочтений пользователя.

ВВЕДЕНИЕ

В наше время разрабатываются различные средства, позволяющие минимизировать человеческие ресурсы, направленные на решение тривиальных задач, обеспечивающие дополнительные ресурсы времени на принятие сложных стратегических решений. Например, на сегодняшний день всё больше людей задумываются над вопросом здорового питания, но большинство не имеют времени или необходимых знаний для обеспечения данного процесса. Поскольку для построения стратегии здорового и сбалансированного питания необходимо обрабатывать много различной информации, о том какой образ жизни ведёт человек, как он питается и каких результатов в использовании энергетических ресурсов организма стремится достичь. Для выполнения этих задач можно разработать средство фиксации достоверной информации о том, как питается пользователь с последующим предоставлением рекомендаций о сбалансированном питании для принятия решений.

I. СИСТЕМА ПРИНЯТИЯ РЕШЕНИЙ ПО УПРАВЛЕНИЮ СБАЛАНСИРОВАННЫМ ПИТАНИЕМ

Для поддержки принятия решений по управлению сбалансированным питанием была осуществлена попытка создания автоматизированной системы, которая:

- предоставляет инструмент сбора информации о составе и графике питания;
- обрабатывает поступающую информацию в соответствии со структурой данных;
- предоставляет релевантные данные по запросам пользователя;
- формирует персональные рекомендации по повышению эффективности питания.

II. ИНСТРУМЕНТ СБОРА ИНФОРМАЦИИ О СОСТАВЕ И ГРАФИКЕ ПИТАНИЯ

Средство фиксации является частью системы, которая представляется следующим образом:

- Наблюдения как происходит процесс. Для реализации необходим отслеживающий механизм.
- Сбор информации. Основной процесс системы, который формирует множество данных для обеспечения функционирования системы.
- Отслеживание изменений в режиме реального времени. Процесс, который на основе текущих данных и сформированной модели принимает решения или выдаёт рекомендации.

Поскольку на текущий момент не существует реальных возможностей использования датчика, который бы обеспечил отслеживания всех необходимых показателей, его можно заменить на смартфон. Используя возможности предоставляемые смартфоном, на его основе можно разработать инструмент фиксации и автоматической оцифровки информации для формирования множества данных. Также он может стать инструментом отслеживания изменений в реальном времени и общения системы и пользователя.

III. ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС ПРИЛОЖЕНИЯ

Приложение для предоставления персональных рекомендаций должно являться частью повседневной жизни пользователя. Таким образом процесс ввода и вывода данных должен быть максимально простым, и быть открытым большому числу пользователей.

Вопрос инфографики является одним из самых актуальных на текущий день. Большинство приложений-конкурентов имеют перегруженный интерфейс, что делает их неудобными и сложными в использовании.

В приложении «HealthAndFood» сделан упор на решение данной проблемы и отображение самой необходимой информации в удобном и простом виде. Одной из ключевых особенностей является возможность работы без постоянного подключения к сети интернет, отложенная

синхронизация данных и получение пакетов обновлений.

Главный экран содержит круговые графики по выполнению норм о воде, килокалориях и главных питательных веществах, подчеркивая результаты следования рекомендациям и мотивируя выполнять нормы.



Рис. 1 – Экраны рекомендаций и истории рациона

Экран поиска блюда сопровождается графиками о процентном составе главных питательных веществ, помогая быстрым образом принять решение о его полезности на текущий момент.



Рис. 2 – Экраны выбора типа приёма пищи и поиска блюд

Экран добавления блюда в рацион содержит графический компонент меры порции в 3 градациях:

- стакан с жидкостью (для жидкостей);
- тарелочка с жидкостью (для супов и жидких блюд);

– тарелочка с гранулами (для гранулированных блюд).



Рис. 3 – Экраны деталей блюда в 3 градациях

Графический компонент меры порции добавляет красивую анимационную составляющую, которая помогает привлечь пользователя для регулярного ввода информации о рационе делая его более интерактивным.

Гибкая система уведомлений позволяет напоминать пользователю если данные о питании не были занесены в журнал. Распределённая система синхронизации базы данных позволяет делиться блюдами и ингредиентами с сообществом с возможностью создавать свои приватные записи.

IV. Выводы

Реализованный мобильным приложением «HealthAndFood» механизм регистрации информации в дневнике питания пользователя превращает смартфон в удобный инструмент, обеспечивающий формирование базы знаний о хронометрии и составе питания индивида. В результате интеллектуального анализа накопленной информации пользователь получает персональные рекомендации сбалансированного питания.

V. СПИСОК ЛИТЕРАТУРЫ

- Recitas, L. The Plan: Eliminate The Suprising Healthy Foods That Are Making You Fat And Lose Weight Fast / L. Recitas// Grand Central Life & Style – 2013. – Vol. 320.
- Герберт, Ш. Жизнь по правилам здоровья. Раздельное питание – основа долголетия / Ш. Герберт // Grand Central Life & Style – 2009. – Vol. 128.
- Holleman, M. iOS Programming: The Big Nerd Ranch Guide / M. Hollemans // Big Nerd Ranch Guides – 2017. – Vol. 416.
- Selander, D. Advanced Apple Debugging & Reverse Engineering / D. Selander // Razeware LLC – 2017. – Vol. 308.
- Todorov, M. Advanced Apple Debugging & Reverse Engineering / M. Todorov // Razeware LLC – 2017. – Vol. 414.
- Львовский, С. М. Набор и вёрстка в системе LaTeX / С. М. Львовский // Издательство: МЦНМО, 2006. – 448 с.

БАЙЕСОВСКИЕ НЕЙРОННЫЕ СЕТИ

Левко С. В., Шадраков Р. Н.

Факультет компьютерных систем и сетей, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {okvel.s, roma8002}@gmail.com

Сегодня в мире существует и используется большое количество цифровых устройств, которые вырабатывают огромное количество информации. Чтобы эта информация приносила пользу необходимо её проанализировать и обработать. Поэтому алгоритмы интеллектуальной обработки данных, использующиеся для этих целей, приобретают все большее значение в современном мире. Байесовские сети являются одним из современных представителей данной группы алгоритмов.

ВВЕДЕНИЕ

В современном мире информация играет огромную роль в жизни людей. Каждый день люди потребляют и производят терабайты информации. На текущий момент важно уметь не только компактно хранить такой объём информации, но и эффективно обрабатывать и анализировать ее. Для этих целей используются алгоритмы интеллектуальной обработки данных. Одним из таких алгоритмов является использование байесовской нейронной сети.

I. ТЕОРЕМА БАЙЕСА

Байесовские сети относятся к группе методов, выделяющих и использующих формализованные закономерности. В данной группе выделяют следующие подгруппы: методы кросстабуляции логической индукции и вывода уравнений. Байесовы сети относятся к методам кросстабуляции. Кросс-табуляция - простая форма анализа данных, где искомое значение находится на пересечении значений атрибутов. В зависимости от количества атрибутов кросстабуляция бывает двумерной и трехмерной.

В основе байесовых сетей лежит теорема Байеса теории вероятностей для определения апостериорных вероятностей составляющих полную группу попарно несовместных событий по их априорным вероятностям. Теорема Байеса -- теорема из области теории вероятностей, с помощью которой можно определить вероятность какого-либо события при условии, что произошло другое событие, взаимозависимое с искомым. Формула (1) является математической записью теоремы Байеса.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}. \quad (1)$$

Для переменных формулы (1) справедливы следующие определения:

- $P(A)$ — априорная вероятность гипотезы A (безусловная вероятность);
- $P(A|B)$ — вероятность гипотезы A при наступлении события B (апостериорная вероятность);

- $P(B|A)$ — вероятность наступления события B при истинности гипотезы A ;
- $P(B)$ — полная вероятность наступления события B [1].

В контексте интеллектуального анализа данных:

- A — параметры модели, которую необходимо обучить;
- B — данные, которые уже имеются;
- $P(A|B)$ — искомое распределение вероятностей параметров модели после того, как приняты во внимание имеющиеся данные;
- $P(B|A)$ — правдоподобие (вероятность данных при условии зафиксированных параметров модели);
- $P(A)$ — математическая формализация того, что было известно о модели до экспериментов.

Таким образом в рамках интеллектуального анализа данных с помощью формулы Байеса можно более точно рассчитать вероятность, используя как ранее известную информацию, так и данные новых наблюдений. Однако для применения теоремы Байеса на практике необходимо оборудование, способное выполнять большое количество вычислений, поэтому данный подход стал активно использоваться только в последнее время.

II. ОСНОВНЫЕ ПРИНЦИПЫ НЕЙРОННОЙ СЕТИ

Искусственная нейронная сеть — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма [2]. Нейронная сеть является самообучающейся системой и имеет большое разнообразие вариантов.

Каждая нейронная сеть состоит из искусственных нейронов, связанных между собой. Искусственный нейрон — это вычислительная единица, которая получает информацию, производит над ней простые вычисления и передает ее дальше. Они делятся на три основных типа: входной, скрытый и выходной. В том случае, ко-

гда нейросеть состоит из большого количества нейронов, вводится термин слой.

У каждого из нейронов есть 2 основных параметра: входные данные и выходные данные. Для входного нейрона, входные и выходные данные одинаковые. Для остальных, в поле входных данных попадает суммарная информация всех нейронов с предыдущего слоя, она нормализуется с помощью функции активации и попадает в поле выходных данных.

Каждый синапс имеет вес, который определяет, насколько соответствующий вход нейрона влияет на его состояние. Состояние нейрона определяется по формуле (2), где n — число входов нейрона, x_i — значение i -го входа нейрона, а w_i — вес i -го синапса.

$$S = \sum_{i=1}^n x_i \cdot w_i. \quad (2)$$

Затем определяется значение аксона нейрона по формуле (3), где f — некоторая функция, которая называется функцией активации.

$$Y = f(S). \quad (3)$$

Функция активации — это способ нормализации входных данных. То есть, если на входе будет большое число, пропустив его через функцию активации, получится число в нужном диапазоне. Существует большое количество функций активации. Основные из них: линейная, сигмоид и гиперболический тангенс. Главные их отличия — это диапазон значений.

Линейная функция почти никогда не используется, за исключением случаев, когда нужно протестировать нейронную сеть или передать значение без преобразований.

Наиболее часто в качестве функции активации используется сигмоид. Его математическая запись представлена формулой (4). Диапазон значений в данной функции от 0 до 1. Основное достоинство этой функции в том, что она дифференцируема на всей оси абсцисс и имеет очень простую производную [3].

$$f(x) = \frac{1}{1 + e^{-ax}}. \quad (4)$$

Гиперболический тангенс используется, только тогда, когда значения могут быть и отрицательными, и положительными, так как диапазон функции от -1 до 1. Использовать эту функцию только с положительными значениями нецелесообразно так как это значительно ухудшит результаты нейросети. Уравнение гиперболического тангенса показано на формуле (5).

$$f(x) = \frac{e^2x - 1}{e^2x + 1}. \quad (5)$$

Синапс — это связь между двумя нейронами. У синапсов есть 1 параметр — вес. Благодаря ему, входная информация изменяется, когда передается от одного нейрона к другому. На вход попадет информация из того нейрона, у которого вес будет больше. Именно благодаря этим весам, входная информация обрабатывается и превращается в результат [22].

III. БАЙЕСОВЫ СЕТИ

Байесовская сеть — нейронная сеть, у которой связь между парой нейронов задана не числом (весом), а вероятностью, высчитываемой по формуле Байеса (см. рис. 1) [4].

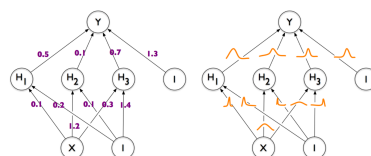


Рис. 1 — Классическая и байесова нейронные сети

Такая система имеет ряд достоинств:

- ситуации, когда значения некоторых переменных неизвестны легко обрабатываются за счёт того, что в модели определяются зависимости между всеми переменными;
- построенные байесовские сети просто интерпретируются и позволяют на этапе прогностического моделирования легко производить анализ по сценарию «что — если»;
- позволяет естественным образом совмещать закономерности, выведенные из данных, и фоновые знания, полученные в явном виде (например, от экспертов);
- позволяет избежать проблемы переподргонки, то есть избыточного усложнения модели, чем страдают многие методы (например, деревья решений и индукция правил) при слишком буквальном следовании распределению зашумленных данных.

Принцип работы байесовской сети такой же как и классической нейронной сети, только вместо веса синапса используется значение вероятности. Такой подход позволяет получать более точные предсказания данных, что повышает ценность метода.

1. Bishop, C. Pattern Recognition and Machine Learning / С. Bishop // Springer. — 2007. — 738 pp.
2. Рашид, Т. Создаём нейронную сеть / Т. Рашид // Вильямс. — 2018. — 272 с.
3. Тулупьев, А. Л. Байесовские сети. Логико-вероятностный подход / А. Л. Тулупьев, С. И. Николенко, А. В. Сироткин // Наука. — 2006. — 608 с.
4. Borgelt, C. Graphical Models: Methods for Data Analysis and Mining / C. Borgelt. // London : Wiley. — 2009. — 404 pp.

ПРОГРАММНОЕ СРЕДСТВО ДЛЯ ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ ПРИ РАЗРАБОТКЕ И ЭКСПЛУАТАЦИИ СИСТЕМЫ АВТОМАТИЗИРОВАННОГО УПРАВЛЕНИЯ ТРАНСПОРТНЫМ ОБОРУДОВАНИЕМ

Марко А. Ф., Чеушев К. В., Лобашинский М. В.

Кафедра высшей математики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: mmmts@bsuir.by

В работе рассматриваются методы и алгоритмы контроля целостности программных продуктов, которые используются в системах с автоматизированным управлением транспортным оборудованием, на этапе их разработки и эксплуатации пользователем.

При разработке и эксплуатации программных продуктов (ПО) важной задачей является обеспечение их цельности, необходимой для предотвращения незапланированных изменений. Это важно для систем управления различного назначения и, особенно, систем управления транспортным оборудованием и мобильными системами различного уровня автоматизации [1].

Контроль за целостностью в предложенном ПО обеспечивается на этапе разработки с помощью внедрения процесса версионирования проектов в интегрированную среду разработки Microsoft Visual Studio (VS) и систему управления версиями Microsoft Team Foundation Server (TFS), а на этапе эксплуатации – через формирования контрольных сумм [2, 3].

В работе были поставлены и решены следующие задачи:

- версионирование проектов в среде VS и системе TFS через разработанное расширение;
- обеспечение контроля целостности ПО на стороне пользователя, с помощью встроенного функционала в программную часть системы.

Для этого в процессе проектирования программного средства было предложено выделить два функциональных модуля. Основные функциональные блоки модулей и их взаимосвязи приведены на рис. 1.

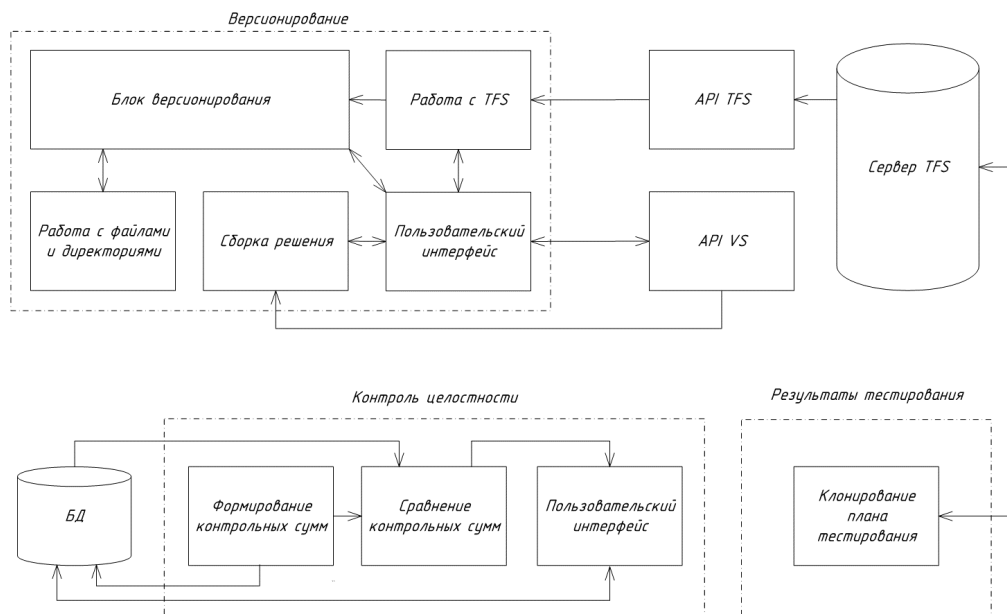


Рис. 1 – Функциональные блоки и взаимосвязи ПО

Основным модулем является модуль версионирования. Он реализован в виде расширения для среды VS. Его пользовательский интерфейс встроен непосредственно в интерфейс VS. Функцией модуля является обновление версий проек-

тов, из которых состоит программная часть системы, при внесении в них изменений. Версионирование возможно в двух основных режимах: в режиме с версионированием только проектов, файлы которых находятся в текущих изменени-

ях и режиме с версионированием всех проектов на основе истории их изменений.

В процессе разработки была реализована концепция [4] формирования версий для проектов, как принадлежащих к версионизируемому решению, так и для проектов, подключённых из других решений по ссылке. В свою очередь проекты, принадлежащие к версионизируемому решению подразделяются на основные проекты, которые являются источником версии последнего релиза и обычные проекты. Определение типа проекта выполняется с помощью структурного анализа файла решения и конфигурационных файлов.

Версия проекта состоит из двух частей: ручной части (первые три старших разряда версии), определяемой последней версией релиза и автоматической части, соответствующей номеру сохранения в системе TFS, в котором был изменён проект. Каждый проект содержит текстовый файл AssemblyInfo, который хранит версию проекта. Для её получения или изменения исполь-

зуются регулярные выражения. В результате после сборки релиза будут сформированы исполняемые файлы, в которых сохранится номер актуальной версии.

Модуль контроля целостности предназначен для определения незапланированных изменений данных на этапе эксплуатации. Для обеспечения целостности данных в процессе эксплуатации на стороне пользователя необходимо учитывать, что некоторые данные не могут оставаться неизменными, поэтому выделяются части, которые должны быть неизменными в процессе эксплуатации, и части, которые могут изменяться.

Программная часть системы состоит из множества различных объектов, таких как исполняемые файлы, файлы данных и объекты баз данных. Формирование хеш-сумм выполняется для каждого типа по-разному. На рис. 2 приведены выделяемые типы и многоступенчатость процесса формирования хеш-сумм для каждого типа.

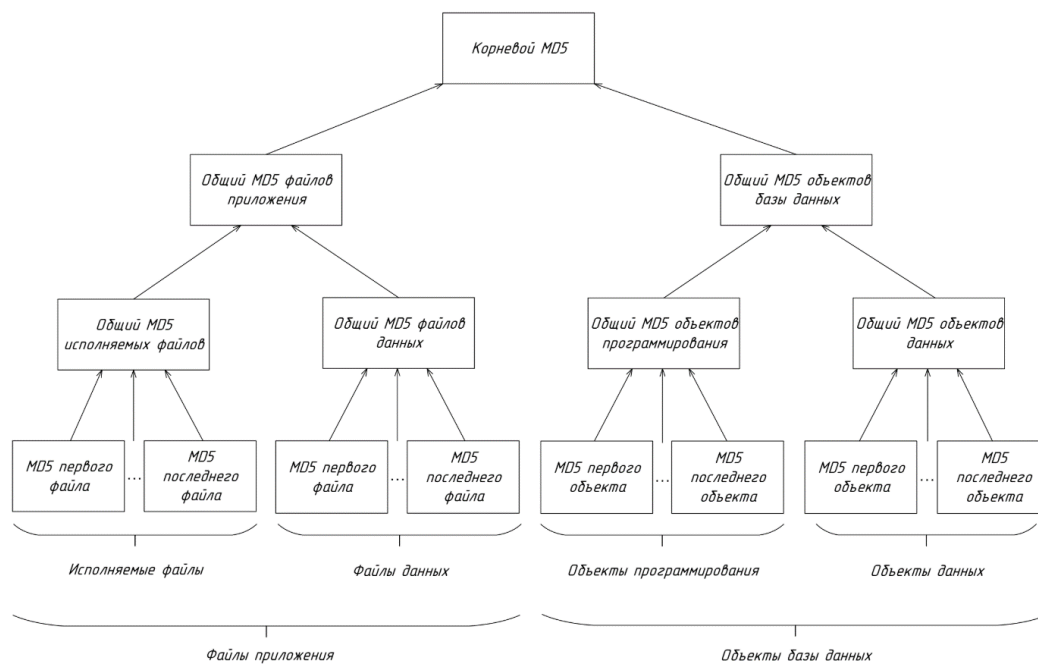


Рис. 2 – Многоступенчатый процесс хеширования

Основная цель многоступенчатого формирования хеш-сумм заключается в предоставлении информации о каждой подгруппе и повышении скорости нахождения изменённых файлов.

Таким образом результатом работы явилось многомодульное программное средство, которое контролирует целостность системы автоматизированного управления транспортным оборудованием в процессе её разработки и эксплуатации на стороне пользователя.

При этом обеспечивается возможность реализации единых подходов к построению алгоритмов компенсации погрешностей координатных систем. К ним относятся погрешности, свя-

занные с изменениями окружающей среды, с локальными неравномерностями шаблонов, с разбросом параметров при изготовлении составных частей координатных систем.

1. Карпович, С.Е. Системы многокоординатных перемещений на механизмах параллельной кинематики. – Минск : Бестпринт, 2017 – 254 с.
2. Arora, T Microsoft Team Foundation Server Cookbook // Т. Arora / Birmingham B3 2PB, UK, 2016 – 309 P.
3. Chowdhury, K Mastering Visual Studio 2017 // К. Chowdhury / Birmingham, UK, 2017 – 433 P.
4. Программирование расширения интегрированной среды разработки VS2017 / А.Ф. Марко, К.В. Чеушев. – ТСЗИ : тез. докл. XVI Белорус.-рос. науч.-техн. конф., Минск, 2018. – С. 63.

ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ НАХОЖДЕНИЯ КОЭФФИЦИЕНТОВ И СТРУКТУРЫ НЕЙРОННОЙ СЕТИ

Нестеренков С. Н., Белов К. П.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: nsn@bsuir.by, showc7@gmail.com

Сформулирована проблема, связанная с обучением нейронной сети. Рассмотрены плюсы и минусы предложенного подхода, а также примеры задач, в которых он может быть применен.

ВВЕДЕНИЕ

Генетический алгоритм - адаптивный поисковый метод, который основан на селекции лучших элементов в популяции, подобно эволюционной теории Ч.Дарвина[1].

Нейронная сеть - это громадный распределенный параллельный процессор, состоящий из элементарных единиц обработки информации, накапливающих экспериментальные знания и предоставляющий их для последующей обработки[2].

Обучение нейронной сети - это, процесс, в котором параметры нейронной сети настраиваются посредством моделирования среды, в которую эта сеть встроена. Тип обучения определяется способом подстройки параметров[2].

Состав нейронной сети:

1. Набор синапсов, каждый из которых характеризуется своим весом или силой.
2. Сумматор - складывает входные сигналы, взвешенные относительно соответствующих синапсов нейрона.
3. Функция активации - ограничивает амплитуду выходного сигнала нейрона.

Для решения задачи при помощи нейронной сети, как правило, приходится самостоятельно обдумывать структуры и тип сети, подбирать функцию активации и начальные весовые коэффициенты, но от этого можно уйти, если перенести процесс подбора полностью на компьютер.

Исходя из описания нейронной сети и генетического алгоритма, можно сделать вывод, что сеть можно обучить, подобрав не только весовые коэффициенты, но и функцию активации, а также топологию.

I. ГЕНЕТИЧЕСКИЙ АЛГОРИТМ

Генетический алгоритм является алгоритмом эволюционных вычислений [3].

Для нахождения экстремума не целочисленной функции в n -мерном пространстве применение графических методов не является возможным (в силу невозможности графической интерпретации n -мерного пространства), а аналитические методы, как правило, требуют помимо исследуемой функции указания дополнительных параметров[4]. Поэтому в n -мерном про-

странстве решение задачи происходит численными методами. Эти методы многомерной оптимизации, в свою очередь, могут быть классифицированы следующим образом:

1. сканирование,
2. покоординатный поиск,
3. градиентные методы
 - 3.1. собственно градиентный,
 - 3.2. метод крутого восхождения,
 - 3.3. метод сопряженных градиентов

Чтобы решить задачу генетическим алгоритмом - необходимо решить задачу оптимизации.

II. ПОСТАНОВКА ЗАДАЧИ ОПТИМИЗАЦИИ

Сформулируем математически задачу оптимизации. Среди элементов x , образующих множества X , необходимо найти такой элемент x^* , который предоставляет минимальное значение $f(x^*)$ заданной функции $f(x)$.

Для корректной постановки задачи оптимизации необходимо задать [1]:

1. Допустимое множество — множество $X = \{\vec{x} | g_i(\vec{x}) \leq 0, i = 1, \dots, m\} \subset \mathbb{R}^n$;
2. Целевую функцию — отображение $X \rightarrow \mathbb{R}$
3. Критерий поиска (max или min).

Тогда решить задачу $f(x) \rightarrow \min_{\vec{x} \in X}$ означает одно из:

Показать, что $X = \emptyset$.

Показать, что целевая функция $f(\vec{x})$ не ограничена снизу.

Найти $\vec{x}^* \in X : f(\vec{x}^*) = \min_{\vec{x} \in X} f(\vec{x})$. Если $\nexists \vec{x}^*$, то найти $\inf_{\vec{x} \in X} f(\vec{x})$.

При решении задачи многомерной оптимизации можно сказать, что \vec{x} является вектором, состоящим из всех элементов векторов целевых функций, значения которых необходимо оптимизировать.

Обучение нейронной сети также можно представить как задачу оптимизации. В таком случае целевой функцией будет являться набор весовых коэффициентов сети. Также возможно осуществить подбор функций активации.

III. ОБЩЕЕ ОПИСАНИЕ РАБОТЫ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Генетический алгоритм состоит из следующих шагов[5]:

1. Случайным образом генерируется конечный набор пробных решений:

$$P^1 = \{p_1^1 \dots p_n^1\}, \quad p_i^1 \in X$$

(первое поколение, n - размер популяции).

2. Оценка приспособленности текущего поколения производится при помощи:

$$F^k = \{f_1^k \dots f_n^k\}, \quad f_i^k = W(p_i^k)$$

в терминологии генетического алгоритма это называется фитнес-функцией.

3. Выход, если выполняется критерий останова (вымирание всех особей или достижение максимально возможного числа итераций, или нахождение решения в пределах доверительного интервала), иначе

4. Генерация нового поколения посредством операторов селекции S , скрещивания C и мутаций M :

$$P^{k+1} = M \cdot C \cdot S(P^k, F^k) \text{ и переход к пункту 2.}$$

В процессе селекции выживают (отбираются) только несколько лучших пробных решений, остальные далее не используются. Скрещивание вместо двух особей создает новую, элементы которой перемешаны каким-то особым образом. Мутация случайным образом меняет случайно выбранную компоненту пробного решения на иную.

IV. РЕШЕНИЕ ЗАДАЧИ С ИСПОЛЬЗОВАНИЕМ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Целевая функция - является компонент, характеризующий весовые коэффициенты, конкретные функции активаций а так же наборы ребер соединяющие нейрон с его соседями.

Пример возможной целевой функции: $F(x) = (x_1, x_2, \dots, x_n)$, где x_1 представляет из себя число - порядковый индекс в списке возможных функций активация, x_2, \dots, x_k - количество связей с нейронами на следующем слое, причем $k \leq l$, где l - количество нейронов на следующем слое.

Фитнес-функция должна быть подобрана таким образом, чтобы показывать, насколько текущие ответы сети соответствуют ожидаемым[5]. Соответственно, для решения конкретной задачи данную функцию необходимо подбирать в зависимости от того, какого результата необходимо добиться.

Шаги решения задачи:

1. Сгенерируем конечный набор пробных решений:

$$P^1 = \{p_1^1 \dots p_n^1\}, \quad p_i^1 \in X,$$

где p_i^1 представляет собой вектор значений компонент целевой функции.

Каждый элемент данного набора представляет собой полноценную нейронную сеть.

2. Произведем оценку приспособленности текущего поколения:

$$F^k = \{f_1^k \dots f_n^k\}, \quad f_i^k = W(p_i^k),$$

где $W(p_i^k)$ - фитнес-функция

3. завершим работу алгоритма, если найдено оптимальное решение, иначе

4. сгенерируем новое поколение, используя функции селекции и скрещивания (каким именно образом следует проводить селекцию и скрещивание, оговаривать не будем, так как подойдет любая из возможных реализаций).

V. ПРИМЕР ИСПОЛЬЗОВАНИЯ ПОДХОДА

Описанный подход обучения может быть применен для нахождения не только весовых коэффициентов нейронной сети, но и для топологии, а так же функций активации. Полученная сеть может быть использована для решения широкого спектра задач в различных областях человеческой деятельности.

VI. ЗАКЛЮЧЕНИЕ

Недостатком использования данного подхода является то, что он требует больше вычислительной мощности, нежели традиционные методы обучения и, соответственно, больше времени, что может быть критичным при большой обучающей выборке. Еще одним недостатком генетического алгоритма является то, что он по сути своей представляет оптимизированный перебор. В связи с этим он может иметь довольно большое, а во многих случаях и неприемлемое время нахождения решения (при условии, что критерием останова является попадание найденного решения в доверительный интервал вероятности, а недостижение порогового значения итераций). В случае, когда критерием останова считается пороговое значение итераций, можно получить решение, не вполне удовлетворяющее ожиданиям.

Достоинством же данного подхода является отсутствие необходимости задания структуры и нахождения оптимальной пороговой функции для самой сети.

VII. СПИСОК ЛИТЕРАТУРЫ

1. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы = Sieci neuronowe, algorytmy genetyczne i systemy rozmyte. — 2-е изд.. 2008. — 452 с. — ISBN 5-93517-103-1.
2. Рутковский Л. Методы и технологии искусственного интеллекта. — М.:, 2010. — 520 с. — ISBN 5-9912-0105-6.
3. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с. — ISBN 978-5-496-02536-2.
4. В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.11
5. Нестеренков, С. Н. Адаптивный поиск вариантов расписания с использованием модифицированного генетического алгоритма / С.Н. Нестеренков // Вести Института современных знаний - 2015. - N2 (63). - С. 67-74

СИСТЕМА ОПРЕДЕЛЕНИЯ ВЕРОЯТНОСТИ УСПЕШНОГО ПРОХОЖДЕНИЯ СОБЕСЕДОВАНИЯ СОИСКАТЕЛЕМ НА ОСНОВЕ ЭКСПЕРТНЫХ ОЦЕНОК

Нестеренков С. Н., Видничук В. Н., Шинкевич Н. Н

кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: nsn@bsuir.by, {vadimxpzl, sn0wf1llin}@gmail.com,

Предложен алгоритм определения вероятности успешного прохождения собеседования соискателем на основе экспертных оценок его навыков, который в дальнейшем будет использован в интеллектуальной системе подбора команды специалистов для решения поставленной задачи.

ВВЕДЕНИЕ

В современном мире всё чаще используются нейронные сети для решения различных задач. Одним из приоритетных направлений является решение задачи классификации. Она заключается в отнесении определённого образца к не пересекающимся множествам. В данной статье предлагается рассмотреть задачу классификации для определения вероятности успешного прохождения собеседования в компании соискателем на основе экспертных оценок. На протяжении некоторого времени собирается статистика о компании, её работниках и в зависимости от этой статистики определяются пространства распределения.

I. СБОР СТАТИСТИКИ

Статистика о компании включает в себя данные о:

- сотрудниках, работающих в данной компании;
- сотрудниках, уволившихся по собственному желанию;
- соискателях, успешно прошедших собеседования в данной компании;
- соискателях, его не прошедших;
- навыках соискателей и сотрудников, требующихся для успешной работы в данной компании.

В статистику соискателей собираются данные с мест его работы, экспертные оценки навыков и времени работы. Экспертные оценки навыков собираются путём прохождения пользователем курсов, его обучения в различных заведениях. Они характеризуют знания соискателя в определённых областях. Каждому навыку ставится своя экспертная оценка. Далее рассчитывается достоверная экспертная оценка, при наличии нескольких факторов, по которым следует оценить каждый из объектов [1-2]. Средняя оценка каждого объекта может быть рассчитана по формуле:

$$m_i = \frac{\sum_{j=1}^m m_{ij}}{\sum_{j=1}^m \sum_{i=1}^n m_{ij}}$$

где m_{ij} – вес i -го объекта, подсчитанный по оценкам всех экспертов.

$$m_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}$$

где x_{ij} – оценка фактора i , данная экспертом j ; n – число факторов; m – число экспертов. Далее выполняется метод простой ранжировки, согласно которому каждый эксперт располагает признаки, основываясь на личных предпочтениях, а затем вычисляется среднее значение важности признака.

Таблица 1 – Таблица ранжировки признаков объекта

	1	2	...	j	m
1	a_{11}	a_{12}	...	a_{1j}	a_{1m}
2	a_{21}	a_{22}	...	a_{2j}	a_{2m}
..
i	a_{i1}	a_{i2}	...	a_{ij}	a_{im}
n	a_{n1}	a_{n2}	...	a_{nj}	a_{nm}

Где a_{ij} – оценка признака экспертом, n – количество признаков, m – количество экспертов.

Данная средняя экспертная оценка используется в различных алгоритмах системы подбора персонала.

Далее собранные данные применяются в качестве своеобразного описания соискателя для последующей обработки и анализа. Целевой переменной является вероятность успешного прохождения собеседования соискателем.

II. ЦЕЛИ И ЗАДАЧИ КЛАССИФИКАЦИИ

Основной целью данной статьи является разработка системы классификации соискателей по группам. Каждая группа представляет собой то, принят ли был на работу соискатель или нет. Так же обязательным является определение вероятности успешного прохождения собеседования соискателем с последующим наймом его на работу. В данном алгоритме каждый образец (соискатель) представлен вектором, включающим в

себя экспертные оценки, идентификатор соискателя, а так же некоторые специфические характеристики данного соискателя, причем последние зависят от конкретной компании и по её желанию могут изменяться. Вышеописанные параметры вектора влияют на принятие решения, а именно к какому из классов относится данный образец. Таким образом классификатор на основе некоторой информации относит объект к одному из классов в связи с определённым распределением пространства.

III. МОДЕЛЬ ДАННЫХ

Для решения поставленной задачи требуются следующие данные: ФИО соискателя и место его работы. Эти данные можно найти в свободном доступе, или же их может предоставить сама компания, например, данные системы подбора персонала. Также требуется информация об экспертных оценках навыков соискателя, которые можно получить по запросу в соответствующие компании. Тут стоит отметить, что не все компании изъявляют желание делиться подобными данными, поэтому с наиболее заинтересованными из них будут заключены соответствующие дополнительные соглашения. Полученные данные будут использованы для обучения нейронной сети с целью классификации соискателей и отнесение их к одной из двух групп: потенциальных работников данной компании или тех, у кого пока недостаточно соответствующих навыков для прохождения собеседования. Таким образом, это поможет ускорить время обработки заявки на подбор персонала в команду, а так же проинформирует соискателя о его сильных сторонах, о навыках, которые необходимо «подтянуть», чтобы получить место в соответствующей компании.

IV. КЛАССИФИКАТОР

Для решения поставленной задачи классификации используются нейронная сеть следующей структуры: один слой Input, на вход которому подаются необработанные данные, и который отвечает за образование вспомогательных параметров для анализа, очистку от строк, содержащих NaN или пустые данные, масштабирование. Далее следует несколько полносвязных слоев. На данном этапе работ структура сети достаточно тривиальна и является по сути многослойным перцептроном. Однако с учетом полученных результатов стоит отметить, что вышеописанную структуру можно усовершенствовать, добавив сверточный слой, который в свою очередь позволит перейти от N-мерного пространства, где расположены параметры каждого вектора характеристик соискателей, к M-мерному, посредством задания соответствующих параметров свертки [3-4]. Активационная функция выходного слоя

нейронной сети представлена на следующем рисунке:

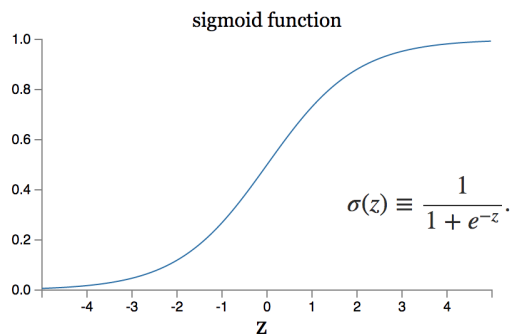


Рис. 1 – Активационная функция выходного слоя нейронной сети

Обновление весов модели происходит по следующей формуле:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t). \quad (1)$$

Где w_{ij} - параметр веса а Δw_{ij} считается по следующей формуле :

$$\Delta w_{ij}(t) = \eta(\delta e_p / \delta w_{ij} + \delta) + \epsilon \Delta w_{ij}(t-1). \quad (2)$$

В качестве примера были сгенерированы данные в размере 100.000 записей. В процессе работы с данными использовался язык программирования Python, пакет обработки данных Pandas, NumPy, пакеты sklearn, keras, matplotlib, mpl_toolkits, причем последние два использовались для построения наглядных изображений разбиения данных на классы [5-6].

1. Нестеренков, С. Н. Математическая модель процесса оценки и подбора персонала на основе многих параметров /С. Н. Нестеренков, В. Н. Видничук // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиозлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2017. - с. 200-201.
2. Нестеренков, С. Н. Генетический алгоритм как способ решения задач многомерной оптимизации / С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиозлектроник ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2017. - с. 196-197.
3. Вандер, Д. П. Python для сложных задач. Наука о данных и машинное обучение /Д. П. Вандер // Питер, 2018. – 576 с.
4. Маккини, У. Python и анализ данных. /У. Маккини // Издательство: ДМК Пресс, 2015. – 482 с.
5. Силен, Д. Основы Data Science и Big Data. Python и наука о данных. /Д. Силен, А. Мейсман, А. Мохамед // Питер, 2017. – 336 с.
6. Machine Learning [Электронный ресурс] / Проект по публикации образовательных материалов в интернете в виде набора онлайн-курсов. – Режим доступа: <https://ru.coursera.org/learn/machine-learning>. – Дата доступа: 16.08.2018.

ОБ АНАЛИТИЧЕСКИХ И ЧИСЛЕННЫХ РЕШЕНИЯХ ДВУХ ТИПОВ РАЗРЕЖЕННЫХ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

Пилипчук Л. А., Лагуто А. А., Пилипчук А. С.

Кафедра компьютерных технологий и систем, кафедра методов оптимального управления

Факультет прикладной математики и информатики

Белорусский государственный университет

Минск, Республика Беларусь

E-mail: pilipchuk@bsu.by, lagutoa@yahoo.com, an.pilipchuk@gmail.com

Рассматривается недоопределённая система, возникающая в задаче об оптимальном расположении сенсоров в узлах графа с целью оценки потоков на ненаблюдаемой части графа. Для двух типов разреженности недоопределённых систем приводятся общие решения. Представлены технологические аспекты построения численных решений разреженных систем указанных типов.

I. ПОСТАНОВКА ЗАДАЧИ

Основой для моделирования процессов оценки потоков на ненаблюдаемой части графа [1] является недоопределённая система линейных алгебраических уравнений вида:

$$\sum_{j \in I_i^+(U)} x_{ij} - \sum_{j \in I_i^-(U)} x_{ji} = \begin{cases} x_i, & i \in I^*, \\ 0, & i \in I \setminus I^*, \end{cases} \quad (1)$$

где $G = (I, U)$ – связный ориентированный двунаправленный граф с множеством узлов I и множеством дуг U , определенных на прямом произведении $I \times I$, $|I| < \infty$, $|U| < \infty$, $I_i^+(U) = \{j : (i, j) \in U\}$, $I_i^-(U) = \{j : (j, i) \in U\}$. I^* – множество узлов с переменной интенсивностью x_i , $I^* \subseteq I$. Двунаправленный граф G обладает следующим свойством: если существует дуга $(i, j) \in U$ с дуговым потоком x_{ij} , то существует и дуга $(j, i) \in U$ с дуговым потоком x_{ji} , который в общем случае может не совпадать с x_{ij} . Для наблюдения за потоками в узлах графа устанавливаются сенсоры. На основе учёта полученной априорной информации от сенсоров система (1) может быть преобразована к следующему виду:

$$\sum_{j \in I_i^+(\bar{U})} x_{ij} - \sum_{j \in I_i^-(\bar{U})} x_{ji} = \begin{cases} x_i + b_i, & i \in \bar{I}^*, \\ b_i, & i \in \bar{I} \setminus \bar{I}^*, \end{cases} \quad (2)$$

$$\sum_{(i,j) \in \bar{U}} \lambda_{ij}^t x_{ij} = 0, \quad t \in T = \{1, \dots, q\}, \quad (3)$$

где b_i , λ_{ij}^t – параметры системы, q – число дополнительных уравнений (3), \bar{I}^* – множество узлов с переменной интенсивностью, $\bar{I}^* \subseteq \bar{I}$. Заметим, что множество T может быть пустым. Граф $\bar{G} = (\bar{I}, \bar{U})$ с множеством узлов \bar{I} и множеством дуг \bar{U} может быть несвязным и/или не двунаправленным. Граф \bar{G} состоит из m связных компонент $\bar{G}_n = (\bar{I}_n, \bar{U}_n)$, $n = \bar{1}, \bar{m}$, \bar{I}_n^* – множество узлов с переменной интенсивностью, $\bar{I}_n^* \subseteq \bar{I}^*$. Определим множества $I_i^+(\bar{U}_n)$, $I_i^-(\bar{U}_n)$ как $I_i^+(\bar{U}_n) = \{j : (i, j) \in \bar{U}_n\}$, $I_i^-(\bar{U}_n) =$

$\{j : (j, i) \in \bar{U}_n\}$. Отметим, что некоторые компоненты связности \bar{G}_n могут не содержать узлов из множества \bar{I}^* , т.е. $\bar{I}_n^* = \emptyset$. Неизвестными системы (2)–(3) являются дуговые потоки x_{ij} , $(i, j) \in \bar{U}$, и переменные интенсивности узлов x_i , $i \in \bar{I}^*$, $\bar{I}^* \subseteq \bar{I}$. Для решения системы (2)–(3) применяются методы декомпозиции разреженной части системы уравнений и её дополнительной части [2]. Матрица разреженной системы (2) имеет блочно-диагональную структуру с различными типами разреженности блоков. Рассмотрим типы разреженности и структуру опоры графа для системы (2) [2,3].

II. ОБЩИЕ РЕШЕНИЯ ДВУХ ТИПОВ РАЗРЕЖЕННЫХ НЕДООПРЕДЕЛЁННЫХ ЛИНЕЙНЫХ СИСТЕМ

Тун 1. Если $\bar{I}_n^* = \emptyset$, то система (2) для компоненты связности \bar{G}_n имеет вид:

$$\sum_{j \in I_i^+(\bar{U}_n)} x_{ij} - \sum_{j \in I_i^-(\bar{U}_n)} x_{ji} = b_i, \quad i \in \bar{I}_n. \quad (4)$$

Опорой $R_n = \bar{U}_R^n$ графа $\bar{G}_n = (\bar{I}_n, \bar{U}_n)$ для системы (4) является покрывающее дерево [2,3]. Общее решение разреженной системы (4) имеет следующий вид [4]:

$$x_{ij} = \sum_{(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n} x_{\tau\rho} \delta_{ij}^{\tau\rho} + \tilde{x}_{ij}, \quad (i, j) \in \bar{U}_n, \quad (5)$$

где $\tilde{x} = (\tilde{x}_{ij}, (i, j) \in \bar{U}_n)$ – некоторое частное решение системы (4), построенное по правилам [5], $\delta(\tau, \rho) = (\delta_{ij}^{\tau\rho}, (i, j) \in \bar{U}_n)$ – характеристический вектор, порожденный дугой $(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n$, относительно покрывающего дерева [5].

Тун 2. Если $\bar{I}_n^* \neq \emptyset$, то система (2) для компоненты связности \bar{G}_n имеет вид:

$$\sum_{j \in I_i^+(\bar{U}_n)} x_{ij} - \sum_{j \in I_i^-(\bar{U}_n)} x_{ji} = \begin{cases} x_i + b_i, & i \in \bar{I}_n^*, \\ b_i, & i \in \bar{I}_n \setminus \bar{I}_n^*. \end{cases} \quad (6)$$

Заметим, что в этом случае опорой графа \bar{G} для системы (2) является лес деревьев, где $\bigcup_{n=1}^m \bar{I}_n = \bar{I}$. Каждое дерево леса содержит ровно один узел из множества опорных узлов $\bar{I}_R^* \neq \emptyset$, $\bar{I}_R^* \subseteq \bar{I}^*$ [2]. Следовательно, опорой графа \bar{G}_n для системы (6) является покрывающее дерево, которое содержит единственный опорный узел с переменной интенсивностью, $|\bar{I}_R^{*n}| = 1$. Общее решение разреженной системы (6) имеет следующий вид [4]:

$$x_{ij} = \sum_{(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n} x_{\tau\rho} \delta_{ij}^{\tau\rho} + \sum_{\gamma \in \bar{I}_n^* \setminus \bar{I}_R^{*n}} x_{\gamma} \delta_{ij}^{\gamma} + \tilde{x}_{ij}, (i, j) \in \bar{U}_R^n, \quad (7)$$

$$x_i = \sum_{(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n} x_{\tau\rho} \delta_i^{\tau\rho} + \sum_{\gamma \in \bar{I}_n^* \setminus \bar{I}_R^{*n}} x_{\gamma} \delta_i^{\gamma} + \tilde{x}_i, i \in \bar{I}_R^{*n}, \quad (8)$$

где вектор $\delta(\tau, \rho) = (\delta_{ij}^{\tau\rho}, (i, j) \in \bar{U}_n; \delta_i^{\tau\rho}, i \in \bar{I}_n^*)$, $(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n$ – характеристический вектор, порождённый дугой $(\tau, \rho) \in \bar{U}_n \setminus \bar{U}_R^n$, относительно покрывающего дерева графа \bar{G}_n ; вектор $\delta(\gamma) = (\delta_{ij}^{\gamma}, (i, j) \in \bar{U}_n; \delta_i^{\gamma}, i \in \bar{I}_n^*)$, $\gamma \in \bar{I}_n^* \setminus \bar{I}_R^{*n}$ – характеристический вектор, порождённый узлом γ , относительно покрывающего дерева графа \bar{G}_n [5].

III. ТЕХНОЛОГИЧЕСКИЕ АСПЕКТЫ ПОСТРОЕНИЯ ЧИСЛЕННЫХ РЕШЕНИЙ РАЗРЕЖЕННЫХ СИСТЕМ

Технологии и алгоритмы построения численных решений системы с типом разреженности 1 (с матрицей инцидентности графа) приведены в [6]. Вычисление компонент характеристических векторов, соответствующих дугам покрывающего дерева, основаны на применении биективного отображения между множеством узлов графа (за исключением одного узла) и множеством дуг покрывающего дерева. В результате применения биекции получено корневое дерево, для хранения которого применяются следующие списковые структуры: список предков узлов ($pred$), список направлений дуг в соответствии с биективным отображением (dir), список династического обхода корневого дерева (d), которое соответствует покрывающему дереву графа, список уровней узлов ($depth$). Корень дерева выбирается произвольным образом.

Технологии и алгоритмы построения численных решений системы с типом разреженности 2 (система полного ранга) основаны на биекции узлов и дуг дерева, поскольку в этом случае число опорных элементов (дуг и узлов) $|\bar{U}_R^n| + |\bar{I}_R^{*n}| = |\bar{U}_R^n| + 1$ совпадает с числом узлов $|\bar{I}_n|$ соответствующего дерева леса для компоненты связности \bar{G}_n . Корнем дерева является единственный узел, входящий в множество \bar{I}_R^{*n} , $|\bar{I}_R^{*n}| = 1$.

Приведём пример базисного графа (рис.1) для системы (2) с двумя компонентами связности $\bar{G}_1 = (\bar{I}_1, \bar{U}_1)$ и $\bar{G}_2 = (\bar{I}_2, \bar{U}_2)$ и типом разреженности 2, $\bar{I}_1^* = \{2, 4, 5\}$, $\bar{I}_2^* = \{6\}$. В результате биективного отображения получен лес корневых деревьев, изображённых на рисунке 2. Списковые структуры хранения леса корневых деревьев приведены в таблице 1.

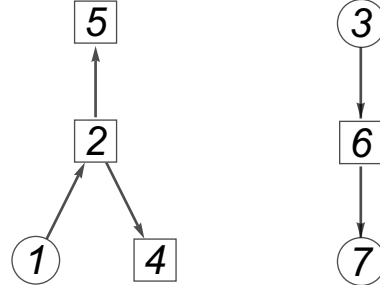


Рис. 1 – Базисный граф для системы (2) с двумя компонентами связности \bar{G}_1 и \bar{G}_2 и типом разреженности 2, $\bar{I}_R^* = \{2, 6\}$

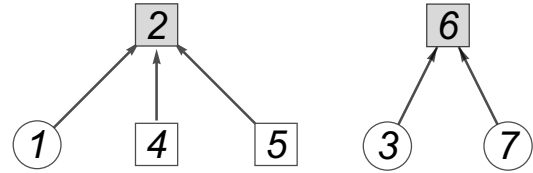


Рис. 2 – Лес корневых деревьев, $\bar{I}_R^{*1} = \{2\}$, $\bar{I}_R^{*2} = \{6\}$

Таблица 1 – Списковые структуры представления леса корневых деревьев для базисного графа системы (2) в случае разреженности типа 2

i	1	2	3	4	5	6	7
$pred[i]$	2	0	6	2	2	0	6
$depth[i]$	1	0	1	1	1	0	1
$dir[i]$	1	0	1	-1	-1	0	-1
$d[i]$	4	1	7	5	2	3	6

1. Пилипчук, А. С. Расположение минимального числа обрезаемых узлов в обобщенном графе для оценки трафика его ненаблюдаемой части / А. С. Пилипчук // Вестник БГУ. Серия 1. Физика. Математика. Информатика. - 2015. - № 1. - с. 108-111.
2. Пилипчук, Л. А. Разреженные недоопределенные системы линейных алгебраических уравнений / Л. А. Пилипчук. – Минск: БГУ, 2012. – 260 с.
3. Габасов, Р. Методы линейного программирования: в 3 ч. / Р. Габасов, Ф. М. Кирилова. Минск: БГУ, 1980. Ч. 3: Специальные задачи. - 368 с.
4. Pilipchuk, L. A. The general solutions of sparse systems with rectangular matrices in the problem of sensors optimal location in the nodes of a generalized graph / L. A. Pilipchuk, O. V. German, A. S. Pilipchuk // Вестник БГУ. Серия 1. Физика. Математика. Информатика. – 2015. – №2. -с. 91-96.
5. Pilipchuk L. A. Sparse Linear Systems and Their Applications / L. A. Pilipchuk - Minsk: BSU, 2013. - 235 p.
6. Пилипчук, Л. А. Алгоритмы и технологии построения решений разреженных недоопределенных линейных систем в Wolfram Mathematica / Л. А. Пилипчук, А. А. Лягуто // Технологии информатизации и управления: сб. науч. ст. в 2 кн. – Минск, 2017. – Вып. 3, кн. 2. – с. 264–272.

О МЕТОДАХ И ТЕХНОЛОГИЯХ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ЭКОЛОГИИ И ВИЗУАЛИЗАЦИИ ПРОСТРАНСТВЕННЫХ ДАННЫХ

Пилипчук А. С., Полячок Е. Н., Пилипчук Л. А., Базаревич О. В., Астапеня П. Н.,
Вишневецкая Т. С.

Факультет прикладной математики и информатики, Кафедра компьютерных технологий и систем,
Белорусский государственный университет
Автотракторный факультет, Кафедра транспортные системы и технологии,
Белорусский национальный технический университет
Минск, Республика Беларусь
E-mail: {an.pilipchuk, arszp10}@gmail.com, pilipchuk@bsu.by, oleg.bazarevich@gmail.com, {astapenia,
vishnevetskaya}@etsconsult.by

Данная работа является продолжением исследований в области моделирования экологических процессов и визуализации пространственных данных на основе систем PostGIS, GeoServer, которые изложены в работе [1]. Разработан программный комплекс мониторинга и оценки выбросов загрязняющих веществ в атмосферный воздух, визуализации мобильных и стационарных источников выбросов. Результатом моделирования и мониторинга является интерактивная карта качества атмосферного воздуха.

ВВЕДЕНИЕ

По сведениям Всемирной организации здравоохранения основным источником ухудшения качества атмосферного воздуха в городах является автомобильный транспорт, оборудованный традиционными двигателями внутреннего сгорания [2]. Например, для города Минска, выбросы загрязняющих веществ в атмосферу от индивидуального автотранспорта составляют около 80%, а общий объем более 85%.

Перспективным направлением повышения качества атмосферного воздуха является использование экологических видов транспорта.

Моделирование экологических процессов и визуализация пространственных данных представляют собой сложные и многофункциональные задачи. Рассматривается задача создания интерактивной карты результатов мониторинга качества атмосферного воздуха. Разработан программный комплекс создания интерактивной карты для хранения, обработки и визуализации геопространственных данных.

I. СТРУКТУРА БАЗЫ ДАННЫХ ДЛЯ СОЗДАНИЯ ГЕОИНФОРМАЦИОННОЙ СИСТЕМЫ

Разработана схема базы данных транспортной сети для хранения результатов замеров. Данные хранятся в двух таблицах:

- одна из которых содержит номера узлов;
- вторая таблица содержит параметры дуг и ссылки на узлы.

Для хранения информации об интенсивности движения транспорта необходимо создать три таблицы, содержащие следующую информацию:

- id узла начала, id узла конца геометрической линии и id направления;
- id узла и параметры замера;

- id замера, тип транспорта, направление и интенсивность.

На основании разработанной схемы базы данных создана картограмма распределения интенсивности движения транспорта на перекрестке, которая представлена на рисунке 1.

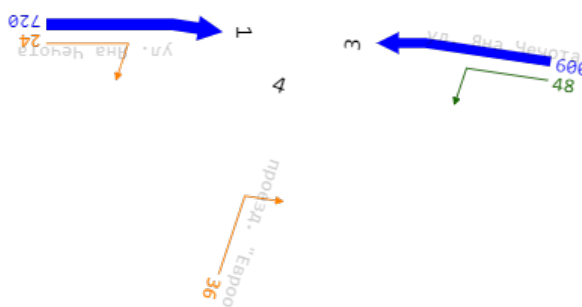


Рис. 1 – Картограмма распределения интенсивности движения транспорта на перекрестке: толщина линии определена в соответствии с агрегацией интенсивности типов транспортных средств

При агрегации данных для формирования средней скорости на дугах транспортной сети используется GPS-треки. На качество GPS-треков влияют такие факторы, как придорожные насаждения, застройка, погрешность устройства и другие.

При обработке GPX файла важным аспектом является фильтрация данных от помех. Скорость движения хранится в отдельной таблице в базе данных. Каждая запись имеет следующую структуру:

- номер дуги;
- дата;
- скорость.

Для ускорения процесса обработки каждого файла его данные загружаются во временную таблицу СУБД PostgreSQL, так как временные таблицы хранятся в ОЗУ.

II. ИНТЕРАКТИВНАЯ ЭКОЛОГИЧЕСКАЯ КАРТА

Для создания интерактивной экологической карты необходимо рассчитать количество загрязняющих веществ, вырабатываемых определённым типом транспортного средства. Объём выбросов загрязняющих веществ в атмосферный воздух при движении транспорта рассчитывается согласно [3]. По причине хранения данных в разных таблицах появляется необходимость их агрегации. Каждый раз, когда пользователь запрашивает данные, система управления базы данных их обрабатывает, но так как они статичны, то их хранение осуществляется в промежуточной таблице в системе управления базы данных PostgreSQL. Для этого были применены материализованные запросы. Для пользователя одним из важнейших критериев работы с интерактивной экологической картой является отзывчивость системы.

На рисунке 2 приведена архитектура программного обеспечения для создания экологической карты.

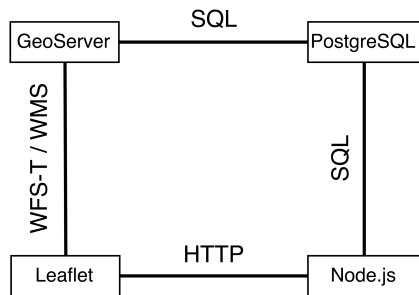


Рис. 2 – Архитектура программного обеспечения

III. ВИЗУАЛИЗАЦИИ ПРОСТРАНСТВЕННЫХ ДАННЫХ

Разработан программный комплекс мониторинга, оценки выбросов загрязняющих веществ в атмосферный воздух и визуализация мобильных и стационарных источников загрязнения.

В результате агрегации входных данных создана таблица с информацией о скорости и интенсивности транспортного потока, которая используется для вычисления объёмов выбросов. На рисунке 3 представлена экологическая карта города с населением свыше 100000 человек.

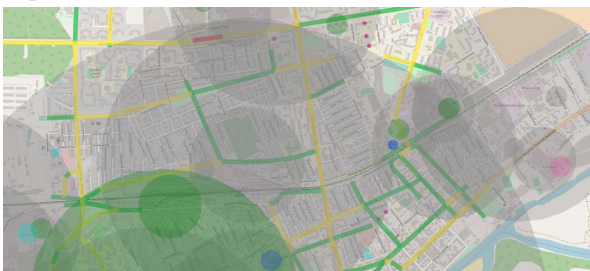


Рис. 3 – Фрагмент экологической карты города с населением свыше 100000 человек

Существенный аспект – иллюстрация просмотра фрагментов экологической карты в интерактивном режиме. Работа с интерактивной

картой позволяет получить информацию об объёмах выбросов загрязняющих веществ для исследуемых объектов. Используя браузер осуществляется редактирование топологии транспортной сети (мобильные источники). Имеется возможность обновления данных о предприятиях (стационарные источники).

Информацию о выбросах в виде диаграммы и табличного представления можно получить, используя инструмент «полигон» и «точка». Выбирая инструмент «полигон» пользователь нажатием на карту ставит точки и замыкает контур. В таблице, полученной после замыкания контура, можно наводить на строку и получать метку центра предприятия. На рисунке 4 приведён пример, иллюстрирующий результат работы инструмента «полигон».

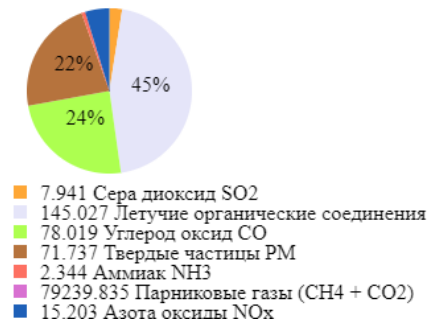


Рис. 4 – Диаграмма суммарных объёмов выбросов загрязняющих веществ

ЗАКЛЮЧЕНИЕ

Разработана структура базы данных для хранения и представления информации о количественных характеристиках загрязнения атмосферного воздуха. Представлена архитектура геоинформационной системы для мониторинга качества воздуха. Разработаны методы и технологии визуализации пространственных данных мониторинга качества атмосферного воздуха.

Разработан программный комплекс создания интерактивной экологической карты. Имеется возможность просмотра растровых слоёв: выбросов загрязняющих веществ, скорости и распределения интенсивности транспортного потока в узлах города.

1. Пилипчук, А. С. Моделирование экологических процессов и визуализация пространственных данных на основе систем PostGIS, GeoServer / А. С. Пилипчук, Е. Н. Полячок, Л. А. Пилипчук, П. Н. Астапеня, Т. С. Вишневецкая // Материалы международной научно-технической конференции «Информационные технологии и системы – ИТС 2017», 25 октября 2017 г. – Минск: БГУИР, 2017. – С. 90–91.
2. Ambient air pollution: A global assessment of exposure and burden of disease [Electronic resource] / World Health Organization. – 2016. – Mode of access: <http://www.who.int/phe/publications/air-pollution-global-assessment/en/>. – Date of access: 24.09.2018.
3. Правила расчета выбросов механическими транспортными средствами в населенных пунктах: ТКП 17.08-03-2006 (02120). – Введ. 28.06.2006. – Минск: Минприроды, 2006. – С. 23.

СРЕДСТВА ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ РАЗРАБОТЧИКОВ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ НА ОСНОВЕ ОТКРЫТОЙ СЕМАНТИЧЕСКОЙ ТЕХНОЛОГИИ

Родионова О. С.

Кафедра интеллектуальных информационных технологий, Факультет информационных технологий и управления, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: oks.rodionova@mail.ru

В данной работе рассматривается необходимость информационной поддержки разработчиков интеллектуальных систем, а также средства информационной поддержки, используемые в рамках Открытой семантической технологии проектирования интеллектуальных систем (OSTIS).

ВВЕДЕНИЕ

В настоящее время на рынке информационных технологий выпускаются всё более комплексные продукты, имеющие сложную внутреннюю структуру и не всегда очевидные внешние связи. К таким продуктам безусловно относятся и интеллектуальные системы. Ввиду указанной сложности разрабатываемых систем в целом и их отдельных компонентов в частности, требуется выделять большое количество ресурсов организации для обучения новых пользователей и разработчиков. Таким образом, время, затрачиваемое на разработку, поддержку и усовершенствование систем значительно увеличивается.

Для решения рассмотренной проблемы предлагается использовать средства информационной поддержки, включаемые в состав разрабатываемой системы или технологии. Указанные средства могут быть оформлены как документация компьютерной системы, руководство пользователя или некоторый справочный компонент самой этой системы в зависимости от её внутреннего устройства.

I. ИНФОРМАЦИОННАЯ ПОДДЕРЖКА В ТЕХНОЛОГИИ OSTIS

В рамках Открытой семантической технологии проектирования интеллектуальных систем[1] в качестве средства информационной поддержки пользователей и разработчиков используется интеллектуальная метасистема IMS[2]. Указанная система содержит знания о моделях, методах и средствах, используемых для проектирования интеллектуальных систем по Технологии OSTIS, ключевых положениях Технологии OSTIS, а также всех прикладных интеллектуальных системах, разрабатываемых с использованием указанной Технологии[3].

Важной особенностью интеллектуальной метасистемы IMS является её динамичность, то есть знания о Технологии, хранимые в рамках системы, обновляются и дополняются вместе с развитием самой этой Технологии, что обеспечивает доступность последних изменений в Техно-

логии как для разработчиков, так и для конечных пользователей.

Другим немаловажным достоинством интеллектуальной метасистемы IMS является то, что знания в ней представлены в формализованном виде, что обеспечивает однозначность понимания представленного материала, лаконичность и быструю масштабируемость накопленных системой знаний. Также формализованное представление позволяет организовать общение системы и конечного пользователя по схеме «вопрос-ответ», в рамках которой пользователь может узнать определение встретившегося понятия, его связи с другими известными системе сущностями, а также область исследования и применения рассматриваемого понятия.

II. НЕДОСТАТОК СУЩЕСТВУЮЩЕГО СРЕДСТВА ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ

Несмотря на рассмотренные выше достоинства системы IMS, ответы на вопросы пользователя не всегда являются исчерпывающими с точки зрения понимания функционирования системы и процессов взаимодействия пользователя с системой.

В качестве примера рассмотрим представленную на рисунке 1 иерархию разделов базы знаний интеллектуальной метасистемы IMS, описывающих различные методики проектирования унифицированных семантических моделей баз знаний. В результате взаимодействия с системой пользователь может получить знания о каждой из представленных методик, их достоинствах, недостатках и аспектах применения.

Однако указанные знания не дают исчерпывающего представления о том, как применять рассмотренные методики построения смысловых моделей баз знаний и реализовывать эти методики с помощью средств и инструментов, используемых в рамках Технологии OSTIS, а также о том, с чего начать знакомство с этими инструментами разработки. Важно отметить, что взаимодействие пользователя и системы по схеме «вопрос-ответ» не предоставляет возможно-

сти для проверки качества усвоенных пользователем знаний.

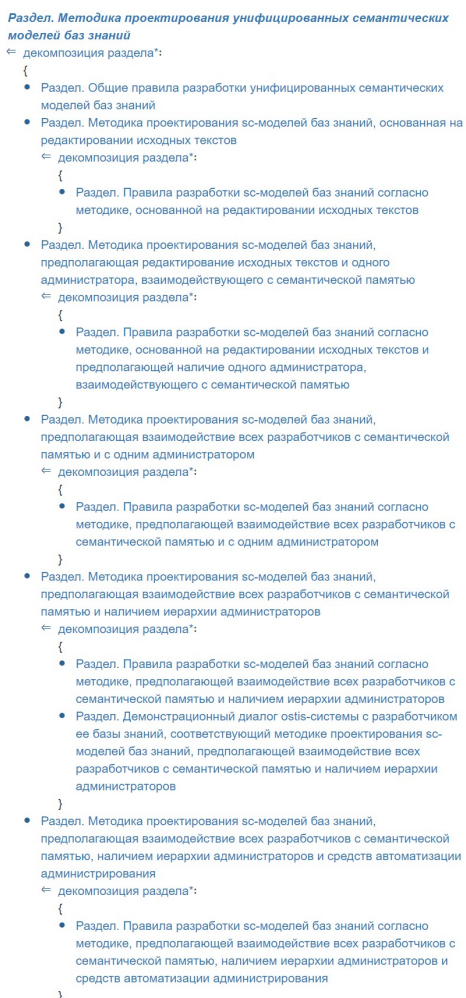


Рис. 1 – Иерархия разделов базы знаний интеллектуальной метасистемы IMS, описывающих различные методики проектирования унифицированных семантических моделей баз знаний

Таким образом, необходимо добавить в средство информационной поддержки пользователей и разработчиков Технологии OSTIS знания о том, как задавать вопросы системе, что означает тот или иной ответ системы, а также сведения о смежных отраслях знаний, непосредственно не освещаемых в рамках самой этой системы.

III. ПРЕДЛАГАЕМОЕ РЕШЕНИЕ

Для расширения возможностей интеллектуальной метасистемы IMS как средства информационной поддержки пользователей и разработчиков систем в рамках Технологии OSTIS было принято решение добавить следующие специальные средства:

- вводные руководства для пользователей и разработчиков;
- упражнения для основных тематических разделов базы знаний интеллектуальной метасистемы IMS;
- формализованное описание типичных ошибок, допускаемых пользователями и разра-

ботчиками при разработке систем, построенных по Технологии OSTIS, и взаимодействию с такими системами;

- формализованное описание часто задаваемых вопросов о Технологии OSTIS в целом и интеллектуальной метасистеме IMS в частности, позволяющее последовательно и методично ознакомиться с указанными Технологией и системой, их возможностями и основными направлениями развития.

Под упражнением к разделу базы знаний интеллектуальной метасистемы IMS понимается некоторое задание, сформулированное на понятном пользователю языке, направленное на проверку усвоения материала, представленного в соответствующем тематическом разделе базы знаний IMS. В качестве задания может выступать формализация некоторого высказывания, определение свойств и связей одного из понятий, рассмотренных в тематическом разделе базы знаний, выделение понятий и предметных областей в некоторой представленной области деятельности. В качестве примера на рисунке 2 представлено упражнение к разделу Раздел. Предметная область параметров и величин. Каждое упражнение содержит ответ для обеспечения возможности проверки результата, полученного пользователем самостоятельно.

Представьте следующее утверждение на языке SСg: "Площадь треугольника равна половине произведения основания треугольника на высоту, проведенную к ней".

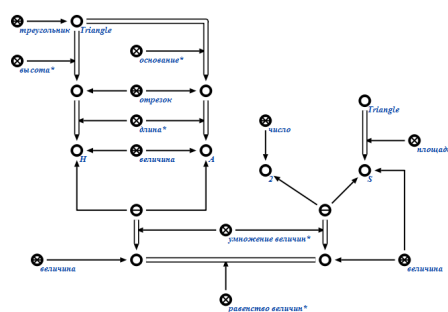


Рис. 2 – Пример упражнения и ответа к нему

ЗАКЛЮЧЕНИЕ

В работе рассмотрено средство информационной поддержки, используемое в рамках Технологии OSTIS, его достоинства и недостатки. Для улучшения качества представленного средства предложено использовать вводные руководства, упражнения, описание типичных ошибок и другие специальные средства.

1. Голенков, В. В., Гулякина, Н. А. Семантическая технология компонентного проектирования систем, управляемых знаниями (OSTIS-2015): Материалы V Междунар. научн.-техн. конф. // Мн.: БГУИР, 2015
2. Метасистема IMS.OSTIS [Электронный ресурс]. – Минск, 2018. – Режим доступа: <http://www.ims.ostis.net>. – Дата доступа: 24.09.2018.
3. Гракова, Н. В., Давыденко, И. Т., Русецкий, К. В. База знаний интеллектуальной метасистемы поддержки проектирования интеллектуальных систем (OSTIS-2014): Материалы IV Междунар. научн.-техн. конф. // Мн.: БГУИР, 2014

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПОИСКА ДОРОЖНЫХ ЗНАКОВ

Шадраков Р. Н., Левко С. В.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {roma8002, okvel.s}@gmail.com

В последнее время автомобильная промышленность совершила большой шаг в создании систем автоматического управления автомобилем. Одним из важнейших компонентов систем автоматического управления автомобилем является система распознавания дорожных знаков. Система распознавание дорожных знаков также может использоваться и отдельно – как система, дающая рекомендации водителю относительно его действий на дороге. Для того, чтобы повысить вероятность правильного распознавания дорожных знаков можно использовать систему на основе сверточной нейронной сети.

ВВЕДЕНИЕ

В настоящее время технологии дистанционного мониторинга, фото- и видеофиксации активно применяются не только в повседневной жизни, но и в решении задач специального назначения. Всё это ведёт к необходимости постоянно поддерживать существующие и развивать новые способы автоматизированной обработки получаемых графических данных. Обработка цифровых изображений часто подразумевает под собой задачу поиска объектов на картинке, в том числе обнаружение, локализацию и распознавание объектов определенного класса.

Для распознавания объекта на изображении необходимо его предварительно локализовать. После этого, область изображения, в которой находится обнаруженный объект, передается классификатору, на выходе у которого будет набор признаков. Используя этот набор можно определить какой именно объект находится на изображении.

Ранее поиск объектов на изображении преимущественно выполнялся при помощи каскадного классификатора с признаками Хаара или каскадных классификаторов на локальных бинарных шаблонах. Однако с развитием сверточных нейронных сетей и их успешным применением в области компьютерного зрения каскадные классификаторы постепенно отходят на второй план.

I. ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ

В случаях, когда необходимо соблюдать последовательность и недостаточно производить распознавание только на одном изображении используются рекуррентные нейронные сети. Эта нейронная сеть представляет собой обычную нейронную сеть, которая имеет обратную связь. Другими словами, на вход нейронной сети или на какой-то из ее слоев передается предыдущее состояние системы. Такой тип нейронных сетей выгодно использовать при работе с потоком изображений. Так как при обработке изображения

из потока могут использоваться данные, которые были получены при обработке предыдущего изображения [2].

Одним из типов сверточных нейронных сетей являются полносверточные нейронные сети. Полносверточные сети – это особый тип искусственных нейронных сетей, результатом работы которых является сегментированное изображение оригинала, где искомые элементы уже выделены требуемым образом. Такие нейронные сети могут использоваться для решения задач, в которых необходимо, например, определить форму и местоположение объекта или нескольких объектов. Самым очевидным отличием полносверточных сетей от других нейросетей является конечный результат работы сети. В результате применения простых сверточных сетей выходными данными являются числа либо массивы чисел [4]. Полносверточные нейронные сети на выходе выдают сегментированное изображение (см. рис. 1). Поэтому их второе название – сегментационные нейронные сети.

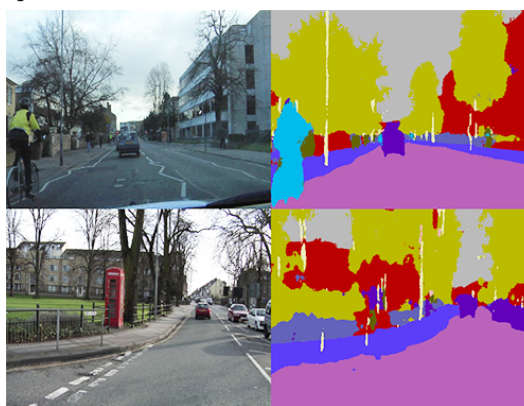


Рис. 1 – Пример работы полносверточной нейронной сети

Полносверточную нейросеть можно использовать, если на изображении могут быть несколько объектов, которые необходимо обнаружить. Следовательно, с ее помощью возможно решить задачу поиска нескольких дорожных знаков на изображении.

II. ПОДГОТОВКА НЕЙРОННОЙ СЕТИ

Так как целью поиска нейронной сети являются только дорожные знаки, ее можно рассматривать как нелинейный фильтр. То есть использование нейросети помогает повысить четкость снимка путем очищения от шума участков внутри определенных областей изображения, не вызывая при этом размывания его краёв.

В выходном изображении, там, где находится объект, значение яркости пикселей равняется единице. В любом другом месте изображения значение яркости пикселей равно нулю. Таким образом, значение яркости пикселей выходного изображения можно считать вероятностью принадлежности данного пикселя к объекту.

Для достижения максимального эффекта при обучении нейронной сети в качестве обучающей выборки использовался набор данных German Traffic Sign Recognition Benchmark (GTSRB), который был использован в соревнованиях по машинному обучению в рамках Международных объединенных конференций по нейронным сетям 2011 [3]. Данный набор данных включает в себя 51 839 изображений, размеры которых колеблются от 15x15 до 250x250 пикселей. Каждое из них содержит один из 43 видов немецких дорожных знаков.

Для того чтобы нейронная сеть могла различать есть на изображении дорожный знак или нет, к GTSRB были добавлены изображения, не содержащие дорожные знаки. В то же время были добавлены изображения, которые имеют большое сходство с искомыми изображениями знаков, например: элементы дорожной разметки, вывески магазинов, фары и колеса автомобилей.

III. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ

Использование нейронных сетей вместо каскадных классификаторов для решения задачи поиска дорожных знаков имеет свои преимущества и недостатки.

Свёрточные нейронные сети также как и каскадные классификаторы с признаками Хаара показывают довольно высокий уровень точности и полноты при поиске объектов на изображении.

Нейронные сети хорошо справляются с различными вариациями наклона объекта, пока каскадные классификаторы неустойчивы к наклонным объектам на изображении. Однако, чтобы обеспечить высокую устойчивость сети в таких ситуациях, необходимо разнообразить обучающую выборку изображениями с подобными преобразованиями.

В то время как каскадные классификаторы с признаками Хаара имеют высокую устойчивость к изменению масштаба объектов на изобра-

жениях, свёрточная нейросеть в некоторых случаях не может распознать объект, при изменении его масштаба.

Для обучения нейронной сети необходимо подготовить большую и качественную обучающую выборку.

ЗАКЛЮЧЕНИЕ

Свёрточные нейронные сети не уступают каскадным классификаторам, а по некоторым параметрам превосходят их, при условии, что на обучение нейросетей будет затрачено достаточно большое количество попыток и масштабы объектов для поиска не будут значительно варьироваться. При использовании сегментационной нейросети можно выделить несколько наиболее важных моментов, которые следует учитывать, чтобы повысить качество ее работы:

- Выборка изображений для обучения должна быть как можно более полной и разнообразной;
- изображения обучающей выборки должны быть наиболее схожими с теми изображениями, для обработки которых будет использоваться сеть;
- при обучении необходимо максимально точно выполнить разметку объектов;
- в процессе создания архитектуры нужно помнить, что универсальной модели, которая способна решить любую задачу, не существует.

Свёрточная нейронная сеть позволяет локализовать объекты быстрее и качественнее, чем каскадные классификаторы, если объект не будет сильно изменяться по масштабу. Эту проблему можно решить путем добавления дополнительной масштабирующей нейронной сети, которая будет предварительно обрабатывать области изображения, в которых возможно наличие дорожных знаков.

1. Рашид, Т. Создаём нейронную сеть / Т. Рашид // Вильямс. – 2018. – 272 с.
2. Галушкин, А. И. Нейронные сети: основы теории // Издательство: «Горячая линия - Телеком», М., 2010. – 480 с.
3. The Institute of Neuroinformatics: Benchmark Website [Electronic resource]. – Mode of access: <http://www.benchmark.ini.rub.de/>. – Date of access: 24.09.2018.
4. Ciresan, U. Meier, J. Masci, L. M. Gambardella, J. Schmidhuber. Flexible, High Performance Convolutional Neural Networks for Image Classification. // Proceedings of the Twenty – Second international joint conference on Artificial Intelligence. – 2011. – P. 1237–1242.
5. Timofte R, Zimmermann K, Van Gool L. Multi-view traffic sign detection, recognition, and 3d localisation. // Workshop on applications of computer vision (WACV). – 2009. – P. 1–8

ОПТИМИЗАЦИЯ ЗАПРОСОВ К БАЗЕ ДАННЫХ С ПОМОЩЬЮ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Шараев Е. В.

Кафедра электронных вычислительных машин, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: eugen.sharaev@gmail.com

В наше время существует большое число решений в области систем управления базами данных. Несмотря на растущую сложность реализации и большое обилие алгоритмов, обеспечивающих большую производительность, по-прежнему существует некоторое число проблем, связанных с производительностью запросов. В то время, как одна часть проблем с производительностью решается с помощью увеличения технических характеристик машин, на которых развернуты базы данных, создания реплик, тонкой настройки конфигураций СУБД либо других оптимизаций, существует так же возможность использования алгоритмов искусственного интеллекта для выбора наиболее оптимальных стратегий для взаимодействия с хранящимися в базе данными.

ВВЕДЕНИЕ

Для минимизации времени выполнения запроса к базе данных в современных РСУБД применяется модуль, который называется планировщик (иногда планировщик/оптимизатор), задачей которого является поиск наиболее оптимального плана выполнения запроса. Планы различаются между собой примененными стратегиями доступа и манипулирования извлекаемыми данными. Примером таких стратегий могут быть *seq scan* (сканирование каждой записи в таблице по очереди), *index only scan* (когда существует покрывающий индекс, при котором в индексе достаточно данных для формирования возвращаемой коллекции и нет необходимости дополнительно обращаться к данным в таблице), различные стратегии объединения – *hash join*, *merge join* и так далее. Таким образом, различия в планах заключаются в различиях использованных алгоритмах в соответствующих планах.

I. СТОИМОСТНАЯ ОПТИМИЗАЦИЯ

В большинстве РСУБД применяются принципы стоимостной оптимизации (CBO, Cost Based Optimizer). При стоимостной оптимизации оптимальность плана запроса определяется значением, которое именуется стоимостью плана. Данное значение рассчитывается как совокупная стоимость каждой операции в запросе и определяется в зависимости от планируемого числа записей, которые будут извлечены из базы (значение *cardinality*). Приблизительное значение числа записей извлекается из статистики, которую собирает и хранит база. При этом отдельно учитывается стоимость для различных операций: последовательное чтение страницы с диска (в PostgreSQL обозначается как *seq_page_cost*), стоимость выполнения операций над кортежами на процессоре (*cpu_tuple_cost*) и так далее. После расчетов стоимости каждого из планов ба-

за сравнивает их и определяет самый производительный на основе полученных значений стоимости.

В некоторых базах так же существует возможность настраивать параметры стоимости различных операций посредством конфигураций, либо даже подсказок планировщику. Данная настройка необходима по причинам того, что база данных может быть развернута на устройствах с различными аппаратными характеристиками. Например, стоимость чтения с диска будет выше, если используется *hdd*, и, соответственно, меньше, если используется *ssd*, стоимость операции на процессоре может быть ниже установленной по умолчанию, если используется процессор нового поколения, либо наоборот – возрастет в случае того, если заведомо известно, что используется непроизводительный процессор.

Таким образом, различие в рассчитанной стоимости каждого из планов обусловлено различием операций, которые производит база в ходе выполнения запроса.

II. ГЕНЕТИЧЕСКИЙ ОПТИМИЗАТОР ЗАПРОСОВ

При большом количестве таких планов время подбора оптимального плана может так же сильно возрастет (так как данная задача относится к классу NP), что может привести к тому, что суммарное время подбора оптимального плана и его выполнение будет неоптимальным. Для предотвращения таких ситуаций в некоторых базах применяются алгоритмы искусственного интеллекта, позволяющие минимизировать суммарное время запроса. Так, например, база данных PostgreSQL использует генетический алгоритм [1] для поиска оптимального плана, когда количество использованных инструкций *join* превышает указанное значение (по умолчанию 12).

Данный алгоритм является более приемлемой альтернативой перебору всех возможных планов при большом их количестве, поскольку при меньшем числе итераций алгоритма (по сравнению с перебором) мы получаем промежуточное решение, которое наиболее вероятно будет являться более оптимальным.

Меньшее количество итераций обеспечивается различиями в критериях останова. При переборе критерием останова является конец цикла по каждому плану, в то время как генетический алгоритм может иметь различные критерии: остановка по прошествии времени либо по достижении указанного числа итераций алгоритма, достижение плато (дальнейшие итерации не дают более оптимальных результатов), а так же комбинации перечисленных.

В данный момент используется именно условие останова по прошествии необходимого числа итераций. Несмотря на свою эффективность, реализация генетического алгоритма в PostgreSQL не является финальной и авторы указывают на то, что все еще необходима работа над поиском оптимальных параметров алгоритма, оптимизация методов скрещивания хромосом и так далее.

III. АДАПТИВНАЯ СТАТИСТИКА В ORACLE DATABASE

Начиная с 12с в Oracle Database был реализован механизм адаптивной статистики [2], который по своему описанию (исходный код данной РСУБД закрыт) представляет собой метод обучения с учителем. Суть данного механизма заключается в оптимизации предсказания количества строк, которые будут извлечены из базы (cardinality). Значение cardinality далее используется при определении оптимальных планов по принципам СВО. После выполнения каждого из запросов база данных сравнивает изначальное значение количества строк, которые планировалось извлечь, и полученное в результате запроса. В случае, когда предсказанное и полученное значения cardinality различаются, база данных делает запись о том, какое число строк вернулось после выполнения и в будущем использует это значение в последующих запросах. Таким образом, сохраняя данные значения, база каждый раз оптимизирует время выполнения.

Так же база обрабатывает случаи потери актуальности статистики. Например, если данные в таблице изменяются больше чем на 10% с момента последнего сбора статистики, то соответствующие значения перестают учитываться при построении планов запросов.

IV. АДАПТИВНАЯ ОПТИМИЗАЦИЯ ЗАПРОСОВ В PostgreSQL

В PostgreSQL так же существует экспериментальная реализация адаптивной оптимизации [3]. Как и в Oracle Database здесь используется аналогичный метод подбора значения cardinality для схожих запросов.

В качестве алгоритма используется модифицированный метод KNN (К-ближайших соседей), способный обрабатывать ситуацию изменения данных в таблице. Хотя данная реализация способна справляться с различными изменениями в данных, авторы все же рекомендуют реинициализировать алгоритм после внесения обширных изменений в данные (в контексте KNN это означает удаление созданных объектов с информацией о cardinality).

Тем не менее, данная реализация имеет несколько недостатков. В качестве признаков здесь используются значения селективности (процент кортежей, удовлетворяющих конкретному условию), извлекаемые из статистики, что может работать не совсем корректно, так как неэквивалентные условия в запросе могут превращаться в эквивалентные по значению результаты предсказания. Например, если положить, что условия $age < 15$ и $age > 25$ имеют одинаковое значение селективности 0.05, то условия $age < 15$ AND reservist IS TRUE и $age > 25$ AND reservist IS TRUE будут иметь эквивалентное предсказание значения cardinality, что не является верным, так как в действительности первый запрос, в отличие от второго, не вернет кортежей (предполагается, что не существует военнообязанных моложе 15 лет). Так же существует проблема с работой данного плагина на репликах, поскольку параметры алгоритма хранятся в таблице в базе (однако существует возможность использования данных с master-сервера).

Альтернативой данному решению может быть использование расширенной статистики [4], которая была реализована в 10 версии PostgreSQL. Она позволяет создавать дополнительную статистику, которая хранит информацию о связи одновременно между несколькими атрибутами в таблице. Основным недостатком данного метода оптимизации является крайне быстро возрастающий размер статистики в зависимости от количества атрибутов, указанных при её создании.

1. PostgreSQL 10.5 Documentation. – 2017 p.
2. Farooq, T. Oracle Exadata Expert's Handbook / T. Farooq, C. Kim, N. Vengurlekar. – 157 p.
3. Ivanov, O. Adaptive Cardinality Estimation / O. Ivanov, S. Bartunov // arXiv:1711.08330. – 2017.
4. PostgreSQL: Documentation: 10: 14.2. Statistics Used by the Planner [Electronic resource]. – Mode of access: <https://www.postgresql.org/docs/10/static/planner-stats.html>. – Date of access: 21.09.2018.

ОЦЕНКА РЕЛЕВАНТНОСТИ КОНТЕНТА В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Шинкевич Н. Н.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: sn0wf1llin@gmail.com

В данной работе рассматривается построение и применение дерева синонимов для определения релевантности контента. В качестве примеров актуальных запросов (актуального контента) выступает информация с Google Trends [1]. Описывается алгоритм нахождения минимального расстояния во взвешенном графе, представляющем текст, которое является мерой релевантности контента примерам запросов Google Trends.

ВВЕДЕНИЕ

Актуальность – один из важных факторов релевантности документа. Если документ содержит устаревшую, неактуальную информацию, его нельзя считать хорошим ответом и не стоит оценивать положительно. Конечно, существуют документы, информация в которых остается актуальной еще долгое время. Например, биография известного художника или писателя. Но многие документы имеют некоторый «срок жизни», в течение которого содержащаяся в них информация актуальна.

I. СОДЕРЖАНИЕ ДОКУМЕНТА

Пусть имеется документ, содержащий некоторую информацию и обладающий определенными признаками, такими как время создания автором и геопозиция. Автора интересует, насколько актуальной является его информация с учетом рассматриваемой темы, а так же перечисленных выше признаков.

Таким образом, необходимо построить систему, позволяющую определить меру релевантности имеющегося документа пользователю актуальным поисковым запросам Google Trends. Мерой релевантности назовем число, в соответствии со значением которого можно проводить сортировку документов по релевантности. Это число должно обладать следующими свойствами:

1. мера релевантности - неотрицательное вещественное число;
2. чем выше релевантность контента, тем выше мера релевантности;
3. мера релевантности должна быть ограничена сверху.

Последнее условие крайне важно с точки зрения удобства пользователя, потому что так он сможет оценить абсолютную степень релевантности контента. Меру релевантности представим в виде целого числа от 0 % до 100 %. Для этого мы округлим и нормируем меру релевантности, подразумевая при этом, что 100 % - релевантный документ представляет собой кон-

тент, содержащий более 20 % слов или более 30 % синонимов запросов самой актуальной выдачи Google Trends.

Для получения актуальных поисковых запросов используются инструменты, называемые web-crawlers (web-spiders), позволяющие собирать и обрабатывать данные автоматически в режиме реального времени. Процесс повторяется каждый час с целью получения самой актуальной информации. В результате работы краулеров в базе данных хранятся записи, каждая из которых представляет собой поисковый запрос в определенный момент времени в том виде, в котором оно представлено в Google Trends.

Каждый полученный поисковый запрос содержит текстовое описание, количество раз ввода поискового запроса в строку поиска Google, график популярности (набор точек, где осью абсцисс представлено время, а ординат - количество раз ввода запроса) и картинку. Вся информация сохраняется в базу данных с целью дальнейшего изучения зависимости популярности контента от степени его соответствия актуальным поисковым запросам в определенный момент времени.

Поисковые запросы разделяются по категориям, каждая категория представлена в Google Trends специфическими для конкретного языка наборами слов по теме поисковых запросов, отсортированных по убыванию количества переходов пользователей. Так же полученные наборы имеют краткие описания (пояснения) к уже приведенным словам.

Реализованная система после поступления запроса пользователя относительно оценки релевантности контента обращается к базе данных, содержащей уже обработанную самую актуальную информацию. Для каждого набора каждой категории строится взвешенный граф, содержащий синонимы данных слов. Для этого используется пакет для анализа естественных языков NLTK Wordnet Synsets [2] на языке программирования Python. Вес ребра графа в данном случае представляет степень близости того или иного синонима. Максимальное число синонимов

для каждого слова было выбрано равным трем, так как выбор большего числа резко увеличивает вычислительную сложность алгоритма. Стоит отметить, что на точность результатов негативно влияет специфика языка, а так же большое количество имен и названий в поисковых запросах. Информация такого рода может не присутствовать в словаре `nlk.wordnet`, или присутствовать с исправлениями, что существенно усложняет поиск вплоть до невозможности его осуществления. В ходе разработки системы было принято решение игнорировать такие слова и не учитывать их в процессах дальнейшего анализа и представления результатов.

Далее при помощи алгоритма Rake [3] из исходного текста извлекаются ключевые слова. Метод Rapid Automatic Keyword Extraction (RAKE) основывается на предположении, что почти всегда ключевыми будут именно те слова, которые содержат значимые слова, но редко включают стоп-слова, местоимения или другие слова с минимальным лексическим значением. Извлечение ключевых слов происходит следующим образом:

1. текст разбивается на слова по позициям стоп-слов и знаков препинания - разделителей;
2. полученные последовательности слов формируют список “кандидатов” в ключевые слова;
3. строится граф встреч данных кандидатов друг с другом в тексте документа;
4. вес каждого слова вычисляется по формуле (1);
5. N слов с наибольшим весом выбираются в качестве ключевых.

$$W(w_i) = \frac{D(w_i)}{F(w_i)}. \quad (1)$$

где (w_i) - некоторое слово, $W(w_i)$ - word weight, $D(w_i)$ - word degree, $F(w_i)$ - word frequency.

После определения ключевых слов для них составляется граф синонимов с глубиной рав-

ной двум. При помощи встроенных алгоритмов `nlk.wordnet` оценивается близость контента к определенному набору слов, представляющему поисковый запрос. Полученное число является усредненным показателем для всех слов, входящих в запрос, по отношению к каждому слову исходного текста.

Тот же процесс повторяется для каждого набора каждой категории, в результате пользователю предоставляется набор поисковых запросов каждой категории, отсортированных по мере релевантности и для удобства просмотра усеченный до трех единиц в каждом наборе.

II. ЗАКЛЮЧЕНИЕ

Таким образом, суть описанного выше подхода к анализу релевантности контента состоит в использовании взвешенного графа синонимов и поисковых запросов Google Trends. Данная методика призвана помочь пользователю в создании актуального контента. Предлагаемый подход успешно применен в рамках практической реализации системы оценки релевантности контента для проекта Exponenta.io, и доступен по ссылке <https://app.exponenta.io>. Дальнейшее развитие применяемой методики может заключаться в разработке системы определения популярности контента на основе вычисленной динамики его соответствия изменяющимся во времени поисковым запросам.

III. СПИСОК ЛИТЕРАТУРЫ

1. Google Trends [Electronic resource] / – Mode of access: https://ru.wikipedia.org/wiki/Google_Trends. – Дата доступа: 08.09.2018.
2. NLTK Wordnet Synsets [Electronic resource] / – Mode of access: <http://www.nltk.org/howto/wordnet.html>. – Date of access: 09.09.2018.
3. Automatic Keyword Extraction from Individual Documents / S. Rose, D. Engel, N. Kramer, W. Cowley // M. W. Berry & J. Kogan (Eds.), Text Mining: Theory and Applications: John Wiley & Sons. – 2010. – Vol.18, № 2. – P. 37–46.

СРЕДСТВА И ПРИМЕРЫ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ ДАННЫХ ГЕОЛОГИЧЕСКИХ МОДЕЛЕЙ

Таранчук В. Б.

Кафедра компьютерных технологий и систем, Факультет прикладной математики и информатики,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: taranchuk@bsu.by

Обсуждаются вопросы использования интегрированного программного комплекса составителя цифровых геологических, геоэкологических моделей. Примерами иллюстрируются предоставляемые системой Mathematica 11 инструменты нейронных сетей для обработки исходных данных цифровых моделей, сопоставления результатов.

ВВЕДЕНИЕ

Моделирование объектов геологии предполагает использование сложных математических методов; программирование, реализация алгоритмов построения и адаптации геологических, геоэкологических моделей трудоемко; для поддержания подобных программных продуктов в актуальном состоянии, отвечающем постоянно возрастающим аппаратным возможностям, нужен коллектив опытных программистов. Учитывая текущие возможности по обеспеченности системами компьютерной математики, эффективным является подход, основанный на интеграции в единый программный комплекс модулей современных версий систем компьютерной алгебры (СКА) и географических информационных систем (ГИС). При этом следует понимать, что для решения задач обработки исходных данных, моделирования никакая конкретная ГИС не является полным набором пространственно-аналитических методов и средств анализа. Во многих случаях приходится комбинировать инструменты ГИС с программами статистического анализа данных, средствами для математически сложных вычислений, алгоритмов анализа и интерпретации пространственных данных.

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ

В [1], [2] описаны предложенные и реализованные методические и технические решения, программные компоненты, которые включены в разработанный интегрированный компьютерный комплекс «Генератор геологической модели залежи» – ГГМЗ. Назначение комплекса тестирование, оценки точности настраиваемых геологических моделей на основе применения СКА, ГИС, «умных» методов адаптации моделей в процессе их эксплуатации, «самонастройки» моделей с учётом дополняемых данных фактического развития процессов. Разработка и функциональное наполнение комплекса ГГМЗ осуществлены с использованием системы компьютерной алгебры Wolfram Mathematica ([3] – [5]) и геоинформационной системы Golden Software Surfer [6]. При

программировании в системе Mathematica модулей графики реализованы технические решения, описанные в [7] – [9]. Предусмотрена возможность, когда программный комплекс в конкретной конфигурации может эксплуатироваться после сборки и сохранения в формате вычисляемых документов CDF [10], [11]. Расчеты и работа пользователя с CDF версией приложения возможны на любом персональном компьютере. При просмотре CDF версии, размещенной на вебсервере, программа просмотра автоматически подгружается в виде плагина браузера. Автономная работа с ПК возможна после установки свободно распространяемого компонента CDF Player. Варианты дополнительных настроек, обеспечивающих интерактивность CDF версии, изложены в [7], [8].

Система Mathematica обеспечивает пользователей широким набором встроенных функций численного анализа [12], методов и функций статистического анализа [13]. Требуемые функции подключаются в ГГМЗ простыми обращениями. Все этапы работы с данными в комплексе включают разнообразные варианты графической визуализации, протоколирование и сопоставление получаемых или помещенных в архив результатов [14], [7], [8]. Инструменты комплекса обеспечивают пользователя возможностями «поиграть» исходными данными и сопоставить результаты с подготовленными эталонами, причем, с возможностями импорта и экспорта данных и изображений, масштабирования. Широкие возможности обмена данными важны при необходимости работы в нескольких программных средах с целью использования уникальных алгоритмов.

II. ИНСТРУМЕНТЫ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Следует понимать, что создание и сопровождение геологической модели не предполагают нахождения единственного решения некоторой математической задачи. Субъективное мнение, квалификация эксперта – факторы, всегда присутствующие в подобной деятельности. При создании моделей приходится оперировать с дан-

ными, точность которых различна, а некоторые исходные данные вовсе противоречивы; на разных участках плотность данных с замерами значительно различается. Поэтому для построения цифровых моделей важно иметь набор инструментов статистической, интеллектуальной обработки данных, коррекции исходной информации.

В комплексе ГГМЗ для визуализации моделей и их сопоставления реализованы более десяти вариантов 1D, 2D и 3D графики, в том числе модули получения и оформления карт и вставок на них, графиков на профилях, 3D визуализации возможных разрезов (вертикальные, горизонтальные), простых и сложных отсечений. Основные возможности программных модулей визуализации для нескольких типичных примеров описаны в упомянутых статьях [1], [2].

В настоящей работе обсуждаются возможности системы Mathematica 11 и поясняются примерами приемы использования и настройки нейронной сети при обработке типичных наборов геологических данных. Иллюстрации даются для комплектов данных, которые извлекаются из формируемых, как в [1] или [2] эталонных распределений. К таковым применяются имеющиеся в комплексе ГГМЗ средства и модули «искажения» эталонной модели – пользователь интерактивно формирует эталонную модель, вносит в нее возмущения. В частности, данные на профилях наблюдений искажаются с использованием функций Mathematica: RandomChoice (случайный выбор), RandomVariate (реализация случайной переменной), SeedRandom (инициализация генератора псевдослучайных чисел), RandomSample (случайная выборка).

Для анализа данных, их интеллектуальной обработки с помощью нейронной сети используются и поясняются варианты настройки функций системы Mathematica 11: NetChain (NetEvaluationMode, TargetDevice, BatchSize, DotPlusLayer, ElementwiseLayer), NetModel (EvaluationFunction, EvaluationNet, SourceMetadata, TrainingSetData, TaskType), NetTrain (Method, BatchSize, MaxTrainingRounds, LearningRateMultipliers, LossFunction, Quantity, ValidationSet, TrainingProgressFunction). Специально подобранными примерами поясняются эффекты изменения параметров сети ValidationSet, Scaled, NetEvaluationMode, BatchSize, TakeDrop.

ЗАКЛЮЧЕНИЕ

В докладе будут изложены и проиллюстрированы примерами возможности системы Mathematica версии 11 по настройке и вариантам применения аппарата искусственных нейронных сетей при анализе исходных данных геологического компьютерного моделирования.

III. СПИСОК ЛИТЕРАТУРЫ

1. Taranchuk, V. B. The integrated computer complex of an estimation and adapting of digital geological models / V. B. Taranchuk // *Studia i Materiały*. Volume 14. – 2017. – № 2. – С. 73–86 (in Engl.)
2. Таранчук, В. Б. Интегрированный программный комплекс тестировщика геологических моделей / В. Б. Таранчук // *Научные ведомости Белгородского государственного университета*. Серия: Экономика. Информатика. – 2017. – № 16 (265). – С. 148–159.
3. Wolfram *MATHEMATICA*. Наиболее полная система для современных технических вычислений в мире [Электронный ресурс] / Режим доступа: <http://www.wolfram.com/mathematica>. – Дата доступа: 23.09.2018.
4. Таранчук, В. Б. Введение в язык Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 51 с.
5. Таранчук, В. Б. Основы программирования на языке Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 49 с.
6. Transform data into knowledge [Электронный ресурс] / Режим доступа: <http://www.goldensoftware.com>. – Дата доступа: 23.09.2018.
7. Таранчук, В. Б. Особенности функционального программирования интерактивных графических приложений / В. Б. Таранчук // *Вестник Самарского государственного университета*. Естественнонаучная серия, раздел Математика: – 2015. – № 6 (128). – С. 178–189.
8. Таранчук, В. Б. Возможности и средства Wolfram Mathematica для разработки интеллектуальных обучающих систем / В. Б. Таранчук // *Научные ведомости Белгородского государственного университета*. Серия: Экономика. Информатика: – 2015. – № 1 (198) том 33. – С. 102–110.
9. Taranchuk, V. B. Methods and tools of development of interactive teaching materials on the computer mechanics / V. B. Taranchuk, M. A. Zhuravkov // *Studia i Materiały*. – 2016. – № 11. – С. 69–83 (in Engl.).
10. CDF. Формат вычисляемых документов – Документы оживают благодаря возможностям вычислений. [Электронный ресурс] / Режим доступа: <http://www.wolfram.com/cdf>. – Дата доступа: 23.09.2018.
11. Таранчук, В. Б. О подготовке и распространении на базе системы Mathematica интерактивных графических приложений / В. Б. Таранчук, В. А. Куликович // *Информатизация образования*: – 2015. – № 1 (75). – С. 3–13.
12. Морозов, А. А. Программирование задач численного анализа в системе Mathematica : учеб. пособие / А. А. Морозов, В. Б. Таранчук // Минск : БГПУ – 2015. – 145 с.
13. Таранчук, В. Б. О возможностях и инструментах реализации вероятностно-статистических методов в системе Mathematica / В. Б. Таранчук // *Теория вероятностей, случайные процессы, математическая статистика и приложения : материалы Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г.А. Медведева*, Минск : РИВШ, 2015. – С. 320–325.
14. Таранчук, В. Б. Графический сервис вычислительного эксперимента : учеб.-метод. пособие / В. Б. Таранчук // Минск : БГУ – 2009. – 124 с.

АВТОМАТИЗИРОВАННЫЙ АНАЛИЗ ВРЕДНОСНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ С ПРИМЕНЕНИЕМ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

Вешторт А. В.

Кафедра интеллектуальных информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ales.veshtort@gmail.com

Статья описывает метод автоматизированного анализа вредоносного программного обеспечения, позволяющий значительно снизить нагрузку на вирусного аналитика, путем предварительной обработки информации о поведении вредоносного образца внутри изолированной среды исполнения приложений.

ВВЕДЕНИЕ

В настоящее время проблема автоматизированного анализа вредоносного программного обеспечения стоит необычайно остро из-за огромных темпов создания новых вредоносных образцов. По данным Virustotal.com [1], только на этот ресурс еженедельно загружается до 400 тыс. образцов вредоносного программного обеспечения.

Очевидно, что ручной анализ такого количества образцов требует больших затрат человеческих и временных ресурсов. Таким образом возникает потребность в автоматизации процесса анализа вредоносного программного обеспечения.

Существующие программные решения лишь частично справляются с задачей автоматизированного анализа вредоносного программного обеспечения, оставляя значительный пласт работы аналитику вредоносных образцов.

Такие решения можно условно разделить на три группы по используемым ими методам анализа вредоносного программного обеспечения: программное обеспечение для сопоставления с образцом; средства статического анализа ПО (интерактивные дизассемблеры); автоматизированные системы динамического анализа ПО («песочницы»).

Сопоставление с образцом является наиболее базовым методом автоматизированного анализа вредоносного программного обеспечения, направленным на обнаружение того подмножества вредоносных образцов, которые обладают набором известных программных артефактов. Таким образом, программное обеспечение, основанное на этом методе анализа, фактически бесполезно при исследовании новых, ранее не исследованных, образцов вредоносного ПО, и может быть использовано только в комбинации с другими средствами анализа.

Существующие средства статического анализа (дизассемблеры) являются программными средствами общего назначения, способными эффективно восстанавливать исходный низкоуровневый код и структуру программного образца,

но не предназначенными конкретно для анализа вредоносных образцов, поэтому задача поиска вредоносного кода перекладывается на плечи аналитика. Кроме того сама сущность статического анализа не позволяет обнаружить изменения кода программного образца в процессе его исполнения.

Программные средства автоматизированного динамического анализа на данный момент представлены системами автоматизированного анализа вредоносного программного обеспечения на основе изолированной среды («песочницы»). Такие системы позволяют получить лишь общую поведенческую информацию об образце: список созданных, удаленных или модифицированных файлов, список созданных ключей реестра Windows, сетевых соединений и т.д. Сама же задача обнаружения характера изменений и исследования переданных данных делегируется аналитику вредоносного программного обеспечения.

I. ПРЕДЛАГАЕМОЕ РЕШЕНИЕ

В данной работе предлагается новый метод автоматизированного анализа вредоносного программного обеспечения, основанный на дополнительной обработке данных, извлекаемых системами динамического анализа на основе изолированной среды.

В частности, предлагается рассматривать список вызовов функций WinAPI как последовательность взаимосвязанных событий, среди которых можно выделить те подпоследовательности, которые соответствуют вредоносному поведению.

Таким образом, задача анализа вредоносного образца сводится к задаче маркировки последовательностей, которая может быть решена с помощью рекуррентной нейронной сети.

II. КОДИРОВАНИЕ ДАННЫХ

Несмотря на то, что категориальные данные в числовом виде могут быть поданы на вход нейронной сети, такой подход может привести к

многочисленным неточностям и ошибкам классификации [2]. Потому предлагается предварительно преобразовать входные данные в one-hot вектора [3].

Для кодирования вызванных функций WinAPI предлагается использовать вектора, составленные из $1 + p$ one-hot векторов, где p - максимальное число параметров функции.

Первый вектор имеет длину t , где t - общее число возвращаемых функцией типов данных; каждому типу данных предварительно присваивается порядковый номер $j, 0 \leq j < t$. При этом если a_i - i -я составляющая вектора, а j - номер возвращаемого функцией типа данных, то $a_i = 1$, если $i = j$, иначе $a_i = 0$.

Все последующие вектора имеют длину $p_t + p_n + 2$ и составлены из двух векторов c и d длиной $p_t + 1$ и $p_n + 1$ соответственно, где p_t - общее число типов данных параметров функции, p_n - общее число имен параметров функции; каждому типу данных параметров функции предварительно присваивается порядковый номер $j, 0 \leq j < p_t$, каждому из имён параметров функции предварительно присваивается порядковый номер $k, 0 \leq k < p_n$.

При этом если c_i - i -я составляющая вектора c , а j - номер типа данных параметра функции, то $c_i = 1$, если $i = j$, иначе $c_i = 0$. Аналогично, если d_i - i -я составляющая вектора d , а k - номер имени параметра функции, то $d_i = 1$, если $i = k$, иначе $d_i = 0$.

Составляющие векторов c и d под номерами p_t и p_n устанавливаются в единицу для «отсутствующих» параметров для функций с общим количеством параметров меньше максимального числа параметров функции. Например, при кодировании функции с p_i параметрами при максимальном числе параметров p_{max} , к функции будет добавлено $p_{max} - p_i$ «отсутствующих» параметров.

III. АРХИТЕКТУРА НЕЙРОННОЙ СЕТИ

Для обработки закодированной последовательности вызовов предлагается использовать двунаправленные нейронные сети, т.к. вызов практически любой функции WinAPI может трактоваться по-разному в зависимости как от предыдущего, так и последующего контекста.

Традиционные рекуррентные нейронные сети обладают крайне ограниченным диапазоном запоминаемых событий. Эта проблема связана с высокой степенью влияния новых входных данных на нейроны скрытого слоя, что приводит к экспоненциальному уменьшению влияния контекстных нейронов предыдущих шагов на вывод сети. Этот эффект называется проблемой исчезающего градиента [4].

Для решения этой проблемы применяются LSTM сети. В таких сетях нейроны скрытого

слоя имеют особую структуру, изображенную на рисунке 1.

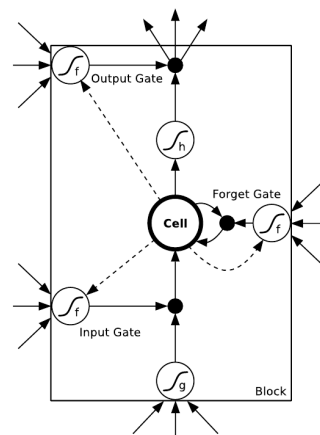


Рис. 1 – Структура LSTM-нейрона

Шлюзы представляют собой нелинейные сумматоры, которые получают сигнал от соседних слоев сети и контролируют активацию нейрона с помощью умножителей (черные точки на рисунке 1). Входной и выходной шлюзы перемножают входные и выходные сигналы нейрона, результат перемножается с предыдущим состоянием нейрона в шлюзе забывания. Функция активации шлюза f , как правило, является сигмоидной. Выходной сигнал нейрона, отправляемый следующему слою сети, формируется умножителем выходного шлюза [4].

Таким образом, в качестве нейронной сети для маркировки последовательностей вызовов была выбрана рекуррентная нейронная сети с одним скрытым LSTM слоем.

IV. ЗАКЛЮЧЕНИЕ

В данной статье был представлен метод автоматизированного анализа вредоносного программного обеспечения, позволяющий значительно уменьшить долю человеческого участия в процессе анализа вредоносного программного обеспечения. В настоящий момент базовый прототип системы, основанной на этом методе, используется в антивирусной лаборатории ОДО «ВирусБлокАда».

1. Virustotal.com - Статистика [Электронный ресурс] Режим доступа: <https://www.virustotal.com/en/statistics/>. - Дата доступа: 31.08.2018.
2. Николаенко, С. И. Глубокое обучение / С. И. Николаенко, А. А. Кадурич, Е. О. Архангельская - СПб.: Питер, 2018 - 480 с.
3. Geron, Aurelien. Hands-On Machine Learning with Scikit-Learn and TensorFlow / Aurelien Geron - 2017. - Vol. 1 - 797 p.
4. Graves, Alex. Supervised Sequence Labelling with Recurrent Neural Networks / Alex Graves - 2010. - Vol. 1 - 137 p.

СОЗДАНИЕ МНОГОВАРИАНТНЫХ ТЕСТОВЫХ ЗАДАНИЙ СРЕДСТВАМИ WOLFRAM MATHEMATICA

Жиленкова В. В.

Механико-математический факультет, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: valeria.zhil@gmail.com

В статье рассматривается вопрос о подготовке многовариантных комплектов зачетных заданий по изучаемым в курсах высшей математики различным темам. Составлены несколько комплектов зачетных заданий на основе фреймов при помощи системы компьютерной алгебры Wolfram Mathematica.

ВВЕДЕНИЕ

Подготовка различающихся комплектов зачетных заданий по различным темам, изучаемым в курсах высшей математики, является актуальной задачей, что особенно важно при реализации учебного процесса, когда надо организовывать текущий контроль знаний нескольких десятков обучаемых. Текущие возможности систем компьютерной алгебры (СКА) позволяют достаточно понятно готовить комплекты с любым числом вариантов заданий одной степени трудности по каждой изучаемой теме. Важно, что использование СКА позволяет давать формулировки упражнений в математической нотации в полном соответствии с обычными учебными материалами. В данной работе программная реализация осуществлена средствами СКА Wolfram Mathematica, в которой по результатам выполнения модуля выводятся два документа в формате PDF: для студентов и для преподавателя. Для каждого студента каждой группы формируются свои индивидуальные варианты заданий (по коду, который включает номер группы и порядковый номер в списке студентов). При подготовке заданий используется генератор случайных целых чисел – именно это обеспечивает различие вариантов. Каждый студент получает свой набор упражнений, для преподавателя выводится отдельный документ с полным перечнем код задания – ответ. Также нужно отметить, что использование СКА позволяет задействовать такую функцию наглядного моделирования как: семантическую функцию, состоящую в расширении знаково-символического опыта оперирования с математическими объектами на основе организации учебной деятельности, построенной на взаимосвязи визуализации и вербализации [1].

Приведем примеры.

I. ЗАДАНИЯ НА РАСКРЫТИЕ НЕОПРЕДЕЛЕННОСТИ «НОЛЬ НА НОЛЬ» ПРИ ВЫЧИСЛЕНИИ ПРЕДЕЛОВ ФУНКЦИЙ

В названном комплекте содержатся два фрейма-примера. Рассматриваются типичные пределы, которые студенты должны уметь идентифицировать и знать алгоритм их решения. Показано, что решающую роль в примерах играют

константы, которые задаются в соответствии с определёнными требованиями.

Пример 1.

$$\lim_{x \rightarrow a} \frac{x^2 - (a+b)x + ab}{k(x-a)} = \quad (1)$$
$$= \frac{(x-b)(x-a)}{k(x-a)} = \frac{a-b}{k}.$$

При этом $a \in (-10, 10)$, $b \in (-12, 18)$, $k \in [-20, 40]$.

Пример 2.

$$\lim_{x \rightarrow \infty} \frac{a_1 x^n + a_2 x^{n-2} + a_3}{b_1 x^m + b_2 x^{m-2} + b_3} = \quad (2)$$
$$= \begin{cases} \frac{a_1}{b_1}, n = m; \\ \ln b, n = \min(k, n, p); \\ 0, n < m. \end{cases}$$

Приведенный предел решается с помощью вынесения общего множителя за скобки, выполнения преобразований, и в зависимости от значений n и m получаются ответы, указанные выше.

II. ИНСТРУМЕНТАРИЙ ПОДГОТОВКИ ШАБЛОНОВ ДИАГНОСТИЧЕСКИХ ЗАДАНИЙ НА ОСНОВЕ ВИЗУАЛИЗАЦИИ

В программном модуле использовались следующие функции системы Wolfram Mathematica: RandomInteger, IntegerPart, Print, Table, Inactive, TraditionalForm, Style, Row [2–3]. Для примера приведем открытую часть кода создания комплекта из двух зачетных заданий для каждого студента группы №9, в которой присутствует три студента (см. рис.1, рис.2).


```

nGroup = 9;
nStud = 3;
Table[
{Print[Style[Row[{"Вариант №", j}], Blue,
18]],
Print[Style[Row[{"Вычислить пределы: "}],
Blue, 18]],
Print[Style[Row[{"1"}], Blue, 20],
Style[TraditionalForm[Inactive[Limit][

$$\frac{x^2 - a[[j]] * x - b[[j]] * x + a[[j]] * b[[j]]}{(k1[[j]] * x - k1[[j]] * a[[j]])}$$
,
x → a[[j]]]], 20]],

```

Рис. 1 – Первая часть программного модуля

В скрытой части кода на основании введенных данных *nGroup* и *nStud* каждому студенту присваивается свой личный и уникальный код, на основании которого формируются все константы, участвующие в формировании каждого задания, задаются на основании данного кода и подчиняются необходимым правилам корректности.

```

Print[Style[Row[{"2"}], Blue, 20],
Style[TraditionalForm[Inactive[Limit][

$$\frac{a1[[j]] * x^{\alpha[[j]]} - a2[[j]] * x^{\alpha[[j]]-1} + a3[[j]]}{b1[[j]] * x^{\beta[[j]]} + b2[[j]] * x^{\beta[[j]]-1} + b3[[j]]}$$
,
x → ∞]], 20]],
Print[
Style[Row[{"n=", α[[j]], "; m=", β[[j]]}],
Blue, 18]],
Print[
Style[Row[{"n=", α[[j]], "; m=", β[[j]]}],
Blue, 18]],
Print[
Style[Row[{"-----"}],
18]], {j, 1, nStud, 1}];

```

Рис. 2 – Вторая часть программного модуля

При запуске модуля программы мы получаем следующий результат (см. рис. 3). При запуске полного модуля программы пользователь получает:

1. PDF файл с наборами заданий «вариант – страница», при этом возможно иное разделение для удобного вывода на печать и последующего использования бумажного варианта заданий;

2. PDF файл с краткими ответами без разделения на страницы (при желании можно отменить вывод данного файла).

Вариант №1

Вычислить пределы:

$$1) \lim_{x \rightarrow -5} \frac{x^2 - 4x - 45}{9.5x + 47.5}$$

$$2) \lim_{x \rightarrow \infty} \frac{5x^{16} + 7x^{15} + 48}{10x^{14} + 4x^{13} - 33}$$

n=16; m=14

Вариант №2

Вычислить пределы:

$$1) \lim_{x \rightarrow 0} \frac{x^2 + 8x}{0. - 4.5x}$$

$$2) \lim_{x \rightarrow \infty} \frac{3x^{14} + 4x^{13} + 28}{12x^{13} + 5x^{12} - 40}$$

n=14; m=13

Вариант №3

Вычислить пределы:

$$1) \lim_{x \rightarrow 3} \frac{x^2 + 5x - 24}{16.5x - 49.5}$$

$$2) \lim_{x \rightarrow \infty} \frac{6x^{12} + 9x^{11} + 58}{13x^{11} + 5x^{10} - 42}$$

n=12; m=11

Рис. 3 – Пример комплекта зачетных заданий

III. СПИСОК ЛИТЕРАТУРЫ

1. Бровка, Н. В. Формы и средства интеграции теории и практики обучения студентов математике / Н. В. Бровка. Минск, БГПУ, 2009. – С. 73–89.
2. Журавков, М. А. Возможности и примеры использования системы Mathematica при преподавании дисциплин и изучении разделов по основам компьютерного моделирования в механике / М. А. Журавков, В. Б. Таранчук // Сетевой журнал «Научный результат». Серия «Информационные технологии», 2016. – Т.1. №1(1). – С. 30–38.
3. Таранчук, В. Б. Основные функции систем компьютерной алгебры / В. Б. Таранчук. Минск, 2015.

Секция «Проектирование
встраиваемых систем»

Section «Embedded systems design»

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ОПТИМИЗАЦИИ ПРЕДСТАВЛЕНИЙ СИСТЕМ БУЛЕВЫХ ФУНКЦИЙ НА ОСНОВЕ БЛОЧНОГО РАЗБИЕНИЯ И РАЗЛОЖЕНИЯ ШЕННОНА

Бибило П. Н., Кириенко Н. А., Ланкевич Ю. Ю.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: kir@newman.bas-net.by

Рассматриваются результаты экспериментального исследования эффективности оптимизации представлений систем булевых функций при синтезе комбинационных схем из библиотечных КМОП-элементов. Процедуры оптимизации включают алгоритмы разбиения систем функций на подсистемы и алгоритмы оптимизации многоуровневых представлений функций на основе разложения Шеннона.

ВВЕДЕНИЕ

Процесс синтеза логических устройств содержит два этапа: технологически независимую оптимизацию и технологическое отображение. На этапе технологического отображения оптимизированные представления описаний логических устройств покрываются описаниями базисных логических элементов заданной библиотеки проектирования.

Эффективными методами технологически независимой оптимизации являются методы, базирующиеся на разложении Шеннона и позволяющие получать оптимизированные BDD (Binary Decision Diagram) [1], а также на разбиении исходного функционального описания систем полностью определенных булевых функций на блоки [2].

В настоящей работе приводятся результаты экспериментального исследования эффективности применения при синтезе логических схем из КМОП-элементов упомянутых выше методов.

I. ОПТИМИЗАЦИЯ МНОГОУРОВНЕВЫХ ПРЕДСТАВЛЕНИЙ, ОСНОВАННАЯ НА РАЗЛОЖЕНИИ ШЕННОНА

Разложением Шеннона полностью определенной булевой функции $f(x)$ по переменной x_i называется представление

$$f(x) = \bar{x}_i f_0 \vee x_i f_1.$$

Подфункции $f_0 = f(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n)$ и $f_1 = f(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n)$ могут быть разложены по одной из переменных $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$. Процесс разложения подфункций заканчивается, когда все n переменных будут использованы для разложения, либо когда все подфункции вырождаются до констант 0, 1.

Под BDDI (Binary Decision Diagram with Inverse Cofactors) в данной работе понимается ориентированный ациклический граф, задающий последовательные разложения Шеннона булевой функции $f(x_1, \dots, x_n)$ по всем ее переменным при заданном порядке (перестановке) переменных, по которым проводятся разложения, с

учетом нахождения инверсных коэффициентов [3]. Возможно построение BDDI двух видов - совместных и отдельных. Совместными BDDI будем называть такие BDDI, которые построены по общей для всех компонентных функций перестановке переменных. Если BDDI для каждой из компонентных функций системы строятся независимо, т.е. так, что каждая из функций разлагается по своей перестановке переменных, то такие BDDI будем называть отдельными.

II. РАЗБИЕНИЕ МНОГОУРОВНЕВЫХ ПРЕДСТАВЛЕНИЙ СИСТЕМ БУЛЕВЫХ ФУНКЦИЙ НА ПОДСИСТЕМЫ

Процедуру разбиения рассмотрим на графовой модели представления логической сети. Ограничениями на сложность каждого из блоков являются число r входных и число q выходных переменных блока. Текущим представлением логической сети является взвешенный ориентированный граф $G = (V, E)$. Вершине $v_i \in V$ соответствует логическое уравнение. Две вершины v_i, v_j графа G связывает дуга e_{ij} , принадлежащая множеству дуг E (дуга исходит из вершины v_i и заходит в вершину v_j), в том и только в том случае, когда выходная переменная уравнения, соответствующего v_i , является входной уравнения, соответствующего v_j .

Задачу разбиения исходной системы логических уравнений на подсистемы можно представить как задачу разбиения множества V вершин графа G на множество непересекающихся подмножеств вершин $\{V_1, \dots, V_p\}$, таких, что $V_k \subset V, V_k \neq \emptyset, V_k \cap V_l = \emptyset, \cup V_k = V$, для $k, l \in \{1, \dots, p\}, k \neq l$. Вершина графа может иметь числовые характеристики, которыми являются числа входных, выходных переменных подсистемы логических уравнений, соответствующих этой вершине.

В настоящем эксперименте исследуется процедура разбиения графа, представляющего систему логических уравнений, на минимальное число p блоков V_k при выполнении заданных ограничений на числа входных (r) и выходных

(q) переменных блоков V_k . В основе процедуры лежит последовательный алгоритм формирования блоков, содержащий этапы выбора начальной вершины для блока, построения множества смежных вершин для блока, последовательного подключения вершин в блок из множества смежных вершин, проверки ограничений на числа входных и выходных переменных. Особенностью алгоритма является выполнение процедуры элиминации внутренних переменных блока с целью их сокращения.

III. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ

Оптимизационные процедуры выполнялись на внутреннем языке SF в экспериментальной системе проектирования логических устройств FLC [4]. С помощью вышеописанных процедур выполнялась оптимизация исходных описаний логических схем в виде системы логических уравнений (формат LOG языка SF). Полученные оптимизированные описания конвертировались в VHDL-описания и подавались на вход синтезатора LeonardoSpectrum, который выполнял синтез схем. Процедура построения совместных BDDI выполнялась программой BDD_Builder [3]. Для блочного разбиения многоуровневых представлений использовалась процедура PartCLTT системы FLC. В качестве примеров использовались многоуровневые описания комбинационной логики из набора Berkeley PLA test set [5]. Результаты эксперимента представлены в таблице 1.

Для каждого примера выполнялся синтез с помощью системы LeonardoSpectrum трех вариантов схем.

Вариант 1. Синтез по исходному описанию комбинационной схемы.

Вариант 2. Оптимизация исходного описания с помощью программы BDD_Builder, синтез по оптимизированному описанию.

Вариант 3. Разбиение исходного описания схемы на блоки с заданными ограничениями на

число входов и выходов, оптимизация каждого блока с помощью программы BDD_Builder, синтез по оптимизированному описанию.

Для каждого тестового примера приведены следующие параметры: числа n и m аргументов и функций системы; величины S_1, S_2, S_3 – площади синтезированных схем по соответствующему варианту, число блоков разбиения p , ограничения на параметры r, q блоков. Площадь схемы оценивается в числе транзисторов.

ЗАКЛЮЧЕНИЕ

Результаты эксперимента позволили установить, что достаточно эффективной процедурой логической оптимизации является оптимизация на основе совместных BDDI. Предварительное разбиение системы функций на подсистемы и последующее построение совместных BDDI позволяет уменьшить площадь схемы (в 8 случаях из 11). При этом уменьшение параметров блоков, на которые разбиваются логические уравнения, позволяет добиваться лучших результатов по площади.

1. Bryant, R. E. Ordered Binary Decision Diagrams / R. E. Bryant, C. Meinel // Logic synthesis and verification /ed. by S. Hassoun, T. Sasao, R. K. Brayton. – [S.1]: Kluwer Academic Publishers. – 2002. – P. 285 – 307.
2. Бибило, П. Н. Оптимизационные преобразования логической схемы на основе блочного разбиения / П. Н. Бибило, Н. А. Кириенко // Информатика. – 2009. – № 3. – С. 5 – 15.
3. Бибило, П. Н. Использование полиномов Жегалкина при минимизации многоуровневых представлений систем булевых функций на основе разложения Шеннона / П. Н. Бибило, Ю. Ю. Ланкевич // Программная инженерия. – 2017. – № 3. – С. 369 – 384.
4. Бибило, П. Н. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов // Минск : Белорусская наука. – 2011. – 279 с.
5. Berkeley PLA test set [Electronic resource] / Mode of access: <http://www1.cs.columbia.edu/cs6861/sis/espresso-examples/>. – Date of access: 29.06.2017.

Таблица 1 – Результаты экспериментального исследования эффективности оптимизации

Пример	n, m	S_1	S_2	S_3	Число блоков p	Параметры r, q
Apex6	135, 94	1830	2110	1858	25	21, 100
C8	28, 18	312	324	318	9	10, 100
Cht	47, 36	680	670	652	7	10, 100
Count	35, 168	256	256	256	8	8, 100
Dalu	75, 16	1834	1396	1286	23	30, 100
x3	135, 99	3462	3788	3154	87	8, 100
term1	34, 10	1044	3086	876	40	8, 100
x4	94, 71	1544	2892	1482	44	10, 100
cordic	23, 2	140	138	138	13	4, 100
ttt2	24, 21	698	766	648	30	10, 100
example2	85, 63	850	996	778	11	15, 100

ИСПОЛЬЗОВАНИЕ СКРИПТОВ ДЛЯ ЛОГИЧЕСКОЙ ОПТИМИЗАЦИИ И СИНТЕЗА ЛОГИЧЕСКИХ СХЕМ

Бибило П. Н., Романов В. И.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: {bibilo, rom}@newman.bas-net.by

Предлагаются программные средства для формирования скриптов, позволяющих выполнять логическое моделирование VHDL-описаний комбинационной логики, формировать соответствующие системы булевых функций, проводить их логическую оптимизацию и синтезировать логические схемы в различных технологических базисах. Пользователю нужно лишь указать исходное алгоритмическое описание, способ логической оптимизации и целевую библиотеку логических элементов, используемую синтезатором LeonardoSpectrum. В результате выполнения скрипта он может выполнить синтез с использованием глобальной оптимизации и сравнить решение, полученное синтезатором LeonardoSpectrum без предварительной оптимизации.

ВВЕДЕНИЕ

Синтезаторы логических схем заменяют каждую конструкцию языка VHDL соответствующим функциональным описанием [1], после чего дальнейшей оптимизации подвергается комбинационная логика, представленная взаимосвязанными логическими выражениями. Такие выражения задают многоуровневые представления систем булевых функций, описывающих функциональные блоки, входящие в состав проекта цифровой схемы, синтезируемой в том или ином базисе логических элементов FPGA (Field-Programmable Gate Array) или ASIC (Application-Specific Integrated Circuits). Используемая логическая оптимизация является, по сути, локальной, т.е. оптимизации подвергаются части схемы (кластеры), выделяемые из оптимизируемого функционального описания проекта схемы. Глобальная оптимизация для достаточно больших проектов не выполняется, так как размерности оптимизационных задач огромны и достигают сотен входных и выходных переменных и сотен тысяч промежуточных логических переменных. В синтезаторе логических схем LeonardoSpectrum [1] размер кластера может быть установлен как управляющая опция синтеза.

Эксперименты [2] показали, что для функциональных описаний комбинационных схем предварительная глобальная оптимизация, выполняемая с помощью программ совместной минимизации систем булевых функций в классе ДНФ (дизъюнктивных нормальных форм) либо с помощью программ оптимизации многоуровневых BDD (Binary Decision Diagram) представлений на основе разложения Шеннона, может давать значительные выигрыши по площади схем и быстродействию по сравнению с результатами синтеза от исходных неоптимизированных функциональных описаний. Однако для применения программ глобальной оптимизации требуется переход от алгоритмических представлений функ-

ций комбинационных блоков ASIC к представлению функций в виде таблицы истинности либо системы ДНФ, которая может быть подвергнута оптимизации с помощью мощных программ, таких как например, ESPRESSO [3], которая позволяет выполнять совместную и раздельную минимизацию систем булевых функций в классе ДНФ по различным критериям либо программа BDD_Builder [4] минимизации многоуровневых BDD-представлений. Если же функции реализуемой системы являются неполностью определенными (частичными), то переход к двухуровневому (табличному) представлению позволяет проводить логическую оптимизацию с учетом возможности доопределения функций. Однако получение таких форм представления систем булевых функций не предусмотрено в синтезаторах, в частности в синтезаторе LeonardoSpectrum, который ориентирован на синтез не только схем FPGA, но и на синтез схем ASIC.

I. ПРЕДЛАГАЕМЫЙ ПОДХОД

Для исходных алгоритмических VHDL-описаний комбинационных блоков ASIC предлагается использовать логическое моделирование, например в системах ModelSim, Questa (ф. Mentor Graphics), для получения VHDL-описаний систем полностью определенных (либо частичных) булевых функций, задающих поведение этих блоков, и предварительную глобальную логическую оптимизацию с помощью программ, имеющихся в системе CMOSLD [5].

Построение эквивалентных описаний для любого алгоритмического VHDL-описания комбинационной логики, выполнение оптимизации и последующий синтез по оптимизированным описаниям осуществляются с помощью разработанной программы, которая автоматически генерирует соответствующий скрипт, пользователю нужно лишь указать исходное алгоритмическое VHDL-описание, способ оптимизации и целевую библиотеку логических элементов, ис-

пользуемую синтезаторе LeonardoSpectrum. В результате выполнения скрипта он может выполнить синтез с использованием глобальной оптимизации и сравнить решение, полученное синтезатором LeonardoSpectrum без предварительной оптимизации. Такой подход к синтезу эффективен для цепочек (конвейеров) арифметических операций с ограниченным числом входных полюсов.

II. ЭТАПЫ ФОРМИРОВАНИЯ СКРИПТА И ВЫПОЛНЕНИЕ ПРОЕКТНЫХ ПРОЦЕДУР

1. По исходному VHDL-описанию комбинационного блока определяется число n входных полюсов данного блока;
2. Генерируется таблица (текстовый файл IN.TST) всех 2^n элементов булева пространства, построенного над переменными булева вектора входных сигналов;
3. Генерируется тестирующая программа для выполнения моделирования исходного алгоритмического описания на всех 2^n наборах значений входных сигналов;
4. Генерируется скрипт для системы VHDL-моделирования, который позволяет выполнить моделирование исходного VHDL-описания комбинационного блока и построить для каждого набора соответствующую реакцию комбинационного блока – так получается текстовый файл OUT.TST реакций комбинационного блока, например:


```
set PRJ tstb
set PRJ_ARCH behavior
set DUT_NAME tstbp0
vlib ./vsim
vmap work ./vsim
vcom -f z:/BAT/vl6/files.f -source
vsim -novopt work.$PRJ +no_glitch_msg
run -a
quit -f
```
5. Вызывается программа формирования (по файлу IN.TST входных наборов и файлу OUT.TST реакций) SF-описания комбинационного блока в виде таблицы истинности;
6. Вызывается одна из программ глобальной логической оптимизации для SF-описания (ESPRESSO, BDD_Builder и др.), в результате получается минимизированное SF-описание системы булевых функций;
7. Вызывается программа конвертации минимизированного SF-описания в VHDL-описание;
8. Генерируется специальный скрипт для синтезатора LeonardoSpectrum, где указываются опции синтеза и требуемая библиотека логических элементов FPGA или ASIC, например:


```
clean_all;
set encoding Gray;
set modgen_select Smallest;
```

```
set asic_auto_dissolve_limit 500;
set auto_dissolve_limit 500;
read z:/BAT/vl6/optim.vhd;
load_library z:/_Mon/libs/s3lib.syn;
set -hierarchy flatten
set effort standard
optimize -target s3lib -macro -area
-effort standard -hierarchy flatten
report_area -cell_usage
report_delay -num_paths 1
-critical_paths
auto_write z:/BAT/vl6/result.vhd
```

9. Вызывается синтезатор LeonardoSpectrum, который выполняет синтез по подготовленному скрипту.

Запуск скрипта позволяет выполнить синтез схемы с предварительной оптимизацией системы булевых функций комбинационного блока ASIC. Возможно формирование двух видов скриптов – для полностью определенных функций и частичных. Ограничения следующие:

- Исходное VHDL-описание комбинационного блока должно содержать только информационные входные и выходные порты типа *std_logic*, *std_logic_vector*;
- Общее число входных полюсов комбинационного блока n не должно превышать 20.

III. ЗАКЛЮЧЕНИЕ

Замена (свертка) цепочек последовательно выполняемых арифметических операций соответствующими матричными моделями полностью определенных либо частичных булевых функций и оптимизация этих моделей является целесообразной при ограниченных размерах задач синтеза, так как позволяет использовать программы глобальной логической оптимизации и получать лучшие результаты при синтезе функциональных комбинационных блоков ASIC.

1. Бибило, П. Н. Системы проектирования интегральных схем на основе языка VHDL. StateCAD, ModelSim, LeonardoSpectrum / П. Н. Бибило // – М.: СОЛОН-Пресс, 2005. – 384 с.
2. Бибило, П. Н. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов / – Минск: Беларус. наука, 2011. – 279 с.
3. Brayton, K. R. Logic minimization algorithm for VLSI synthesis / K. R. Brayton, G. D. Hachtel, C. T. McMullen, A. L. Sangiovanni-Vincentelli // Boston, e.a.: Kluwer Academic Publishers, 1984. – 193 p.
4. Бибило, П. Н. Использование полиномов Жегалкина при минимизации многоуровневых представлений систем булевых функций на основе разложения Шеннона / П. Н. Бибило, Ю. Ю. Ланкевич // Программная инженерия. – 2017. – № 3. – С. 369–384.
5. Бибило, П. Н. Система логического проектирования функциональных блоков заказных КМОП СВИС с пониженным энергопотреблением / П. Н. Бибило [и др.]. // Микроэлектроника. – 2018. – Т. 46. – № 1. – С. 72–88.

РЕШЕНИЕ ЗАДАЧИ СИНТЕЗА ЧАСТОТНО-ИЗБИРАТЕЛЬНОЙ ЦЕПИ С ОПТИМАЛЬНЫМИ ХАРАКТЕРИСТИКАМИ С ПОМОЩЬЮ АЛГОРИТМА ОПТИМАЛЬНОГО СИНТЕЗА ПО ЗАДАННОМУ КРИТЕРИЮ

Дубовик И. А.

Кафедра автоматики, радиолокации и приёмо-передающих устройств, Военная академия Республика Беларусь
Минск, Республика Беларусь
E-mail: duba-77@mail.ru

В данной докладе показан пример решения задачи синтеза частотно-избирательной цепи с оптимальными характеристиками с помощью алгоритма оптимального синтеза по заданному критерию

ВВЕДЕНИЕ

Проектирование устройств с оптимальными частотными характеристиками является одной из важнейших радиотехнических проблем. Чаще всего требуемая реакция цепи задаётся в форме графика или таблицы, а наиболее подходящую функцию цепи необходимо получить по заданной идеальной характеристике [1]. Для решения данной задачи используются аппроксимирующие функции (АФ).

По полученным результатам в [2], можно сделать вывод о том, что для улучшения одного параметра АФ зачастую необходимо ухудшить другие. Это обусловлено низкими вариативными свойствами аппроксимирующей функций. Из-за этого разработчику приходится отказываться от выбранной АФ или увеличивать ее порядок, что приводит к усложнению реализации функции. Таким образом, проблема аппроксимации является первым шагом в синтезе частотно-избирательных цепей.

СИНТЕЗ ЧАСТОТНО-ИЗБИРАТЕЛЬНОЙ ЦЕПИ С ПОМОЩЬЮ АЛГОРИТМА ОПТИМАЛЬНОГО СИНТЕЗА ПО ЗАДАННОМУ КРИТЕРИЮ

Частотно-избирательную цепь, возможно синтезировать различными способами. Это обстоятельство предоставляет возможность использовать, для выбора из совокупности вариантов, наилучший по некоторому признаку, что позволит осуществить оптимальный синтез [3]. Существует множество параметров частотно-избирательной цепи, которые целесообразно оптимизировать. Однако совершенно очевидно, что принцип оптимизации неизбежно должен присутствовать в самой постановке проблемы синтеза, т.е. решение задачи аппроксимации. Только при этом условии можно осуществить оптимальный синтез цепи.

Для решение задачи аппроксимации был разработан алгоритм (рисунок 1). Данный алгоритм позволяет осуществить оптимальный синтез частотно-избирательной цепи за счет вы-

бор наиболее подходящей АФ, которая наилучшим образом воспроизводит требуемые характеристики цепи по отношению к другим АФ на основании критерия, по которому оптимизируется частотно-избирательная цепь. В качестве критериев в алгоритме используются такие критерии как: минимизация порядка частотно-избирательной цепи, минимизация величины неравномерности ГВЗ и минимизация ошибки аппроксимации при заданных значениях параметров АФ.

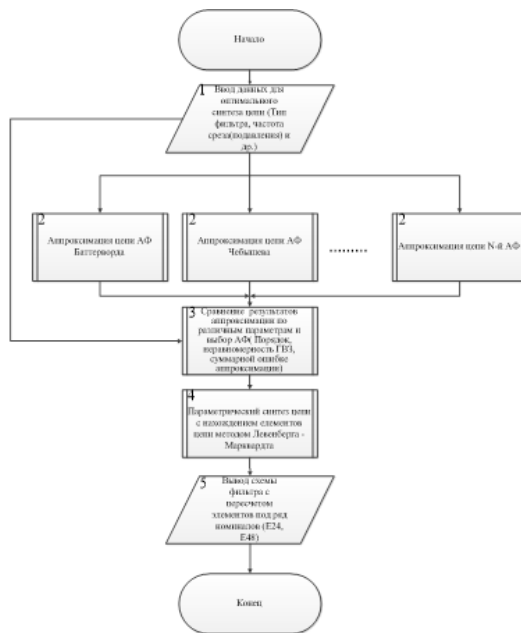


Рис. 1 – Алгоритм оптимального синтеза частотно-избирательной по заданному критерию

Работа алгоритма начинается с ввода необходимых параметров для его работы (тип фильтра (НЧ, ВЧ, ПП, ПЗ), частота среза, частота подавления, искажение в ПП, коэффициент подавления, граничные частоты и (для ПП или ПЗ типов частотно – избирательных цепей), критерий или критерии по которым будет осуществляться оптимальный синтез, доп. параметры для АФ). После ввода входных параметров осуществляется аппроксимация требуемой характеристики.

Определяется порядок, полюса АФ, для определения расположения полюсов используется аналитическое продолжение на плоскость комплексной частоты $s = \pm\sigma \pm j\omega$. После преобразования функции передачи, на плоскость комплексную частоты, находится функция коэффициента отражения из выражения (1) и вычисляются полюса и нули функции коэффициента отражения.

$$K(-s^2) = 1 - \rho(s)\rho(-s) \quad (1)$$

Используются полюса находящиеся в левой полуплоскости и нули передачи АФ находящиеся на мнимой оси комплексной плоскости. Нули передачи АФ не должны быть кратными для получения устойчивости и физически реализуемости частотно-избирательной цепи. После вычисляется функция входного сопротивления НЧП и приводится к виду выражению (2):

$$Z(s) = \frac{\alpha_0 + \alpha_1 s + \dots + \alpha_{m-1} s^{m-1}}{\beta_0 + \beta_1 s + \dots + \beta_m s^m} \quad (2)$$

для последующего вычисления КПМ (коэффициента передачи мощности) по формуле, показанной в выражении (3):

$$K_p(\omega) = \frac{4R_c Re(Z(\omega))}{(Re(Z(\omega)) + 1)^2 + Im(Z(\omega))^2} \quad (3)$$

где n – степень полинома числителя; m – степень полинома знаменателя; $Re(Z(\omega))$ – реальная часть функции входного сопротивления; $Im(Z(\omega))$ – мнимая часть функции входного сопротивления.

После аппроксимации требуемой характеристики, в зависимости от критерия (критериев) оптимизации, которые были определены на первом этапе работы алгоритма, будет выбрана одна из АФ, которая наилучшим образом воспроизводит требуемую характеристику по заданному критерию или критериям.

В случае выбора оптимального критерия "оптимизация по наименьшему порядку" рассчитанные значения порядка АФ сравниваются между собой и выбирается та АФ, которая обеспечивает наименьший порядок. Если в качестве оптимального критерия была выбрана "оптимизация по неравномерности ГВЗ" то тогда вычисляется функция ГВЗ АФ по формуле (4):

$$T(\omega) = -\partial(\alpha \tan(Im(Z(\omega))/Re(Z(\omega))))/\partial\omega \quad (4)$$

В последующей находится неравномерность функции ГВЗ в полосе $[0, \omega]$. После вычисляется максимальное и минимальное значение функции ГВЗ в этой полосе с последующим вычислением разности максимального и минимального значения по формуле (5):

$$\max(T(\omega)) - \min(T(\omega)) = \theta \quad (5)$$

где θ – неравномерность ГВЗ функции входного сопротивления частотно – избирательной цепи.

Выбирается та АФ, которая имеет минимальное значение θ по отношению к другим АФ. В случае выбора оптимального критерия "величина суммарной квадратичной ошибки аппроксимации оценивается ошибка аппроксимации по среднеквадратичному критерию в полосе пропускания и в полосе подавления.

Если выбрано несколько критериев оптимального синтеза порядок выбора осуществляется в той последовательности, который был указан на первом этапе работы алгоритма.

Выбрав АФ, которая наилучшим образом воспроизводит требуемую характеристику по отношению к другим АФ, осуществляется построение лестничной схемы с количеством плеч равных порядку функции сопротивления частотно-избирательной цепи. Для расчета элементов цепи используется метод неопределенных коэффициентов, который был описан в [2], с нахождением значения элементов цепи методом Левенберга – Марквардта. В зависимости от типа частотно-избирательной цепи (НЧ, ВЧ, ПП, ПЗ), которые были определены на первом этапе работы алгоритма, осуществляется преобразование НЧП в один из типов частотно-избирательной цепи.

Конечным итогом работы алгоритма является приведение значений элементов к элементам под ряд номиналов (Е24, Е48). В качестве проверки работоспособности алгоритма был синтезирован полосовой фильтр с параметрами: несущая частота 10МГц, полоса 4МГц, коэффициент подавления 60Дб, величина искажения 0,5 Дб, критерий: минимизация неравномерности ГВЗ. По итогам синтеза частотно-избирательной цепи была выбрана АФ Баттерворда [2]. Данная АФ обеспечивает минимальное значение неравномерности ГВЗ по отношению к другим АФ. Полученные результаты не противоречат апробированным данным [1].

ЗАКЛЮЧЕНИЕ

Благодаря использованию алгоритма оптимального синтеза частотно-избирательной цепи по заданному критерию решается задача аппроксимации, которая является первым шагом в синтезе частотно-избирательных цепей. Решение данной задачи позволяет наиболее точно воспроизвести требуемую характеристику цепи. Таким образом, потребитель получит частотно-избирательной цепи с требуемой оптимальной характеристикой.

СПИСОК ЛИТЕРАТУРЫ

1. Карни, Ш., Теория цепей. Анализ и синтез/Ш.Карни. – М. Связь, 1973. – 269 с.
2. Дубовик, И.А. Способы аппроксимаций частотных характеристик в задачах широкополосного согласования и фильтрации: дис. маг-ра тех. наук: 1-39 80 02 / И.А. Дубовик. – Минск, 2018. – 92 л.
3. Ланнэ, А.А. Оптимальный синтез линейных электрических цепей / А.А. Ланнэ. – М.: Связь, 1969. – 294 с.

СИСТЕМА СБОРА И АНАЛИЗА ДАННЫХ ДЛЯ ПРИЛОЖЕНИЙ РЕАЛЬНОГО ВРЕМЕНИ

Емельянов Д. В.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: dimaplusplus@gmail.com

Широкое распространение получили устройства относящиеся к концепции Internet of Things, в частности умные счетчики. Предполагается, что в будущем устройства станут активными участниками информационных и социальных процессов, где они смогут взаимодействовать и общаться с окружающей средой. Современное общество нуждается в более экономном использовании ресурсов, в том числе электроэнергии, чего позволяют добиться умные счетчики.

ВВЕДЕНИЕ

Все большую популярность приобретают системы сбора и анализа данных в реальном времени, так называемые умные устройства, которые должны уметь передавать интересующую информацию на другое устройство, компьютер или мобильный телефон.

Умные устройства позволяют анализировать информацию в зависимости от поставленных целей. Хранимая и передаваемая информация может носить различный характер, например, температура в исследовательской лаборатории, утечки на нефтеперерабатывающих заводах, ритм сердца. В рамках данного исследования мы сосредоточили свое внимание на измерении электроэнергии.

На примере умного счетчика электроэнергии такая система позволяет реализовать многотарифный учет электрической энергии, иметь более точные и различные показатели измерения, что выгодно и для потребителей, и для коммунальных служб, и для общества в целом.

В результате исследования разработано умное устройство сбора и анализа, позволяющее вести учет в различных аспектах энергопотребления. Разработанное устройство представлено на Рисунке 1.



Рис. 1 – Разработанный счетчик электроэнергии

I. БЕСПРОВОДНЫЕ ТЕХНОЛОГИИ ПЕРЕДАЧИ ДАННЫХ

В рамках нашего исследования нас интересует связь на короткие расстояния для снятия показаний со счетчиков. Рассмотрим наиболее востребованные стандарты беспроводной связи: BLE, ZigBee и Wi-Fi.

Стандарты регламентируют физический и канальный уровни модели OSI, транспортный и сетевой уровень приложения разрабатывается производителями чипов. Для разработчика остается только написать верхний уровень приложения взаимодействующий с устройствами в сети и реализующий бизнес-логику [1].

Таблица 1 – Сравнительная таблица беспроводных технологий передачи данных

	BLE	ZigBee	Wi-Fi 802.11a/b
Частота передачи	2.4 ГГц	868/915 МГц; 2.4 ГГц	5 ГГц
Скорость передачи	1 Мбит	250 Кбит	54 Мбит
Дальность передачи, м	10	10-100	100
Среднее пиковое потребление, мА	50	40	220
Среднее время автономной работы устройства, дней	1-7	более 100	0.5 - 2

Wi-Fi предоставляет высокую скорость и дальность передачи, в то время как Bluetooth и ZigBee имеют гораздо меньшие показатели как в дальности, так и в пропускной способности. Но, стоит отметить, что главное преимущество Bluetooth и ZigBee - энергопотребление. Однако для счетчика электроэнергии время автономной работы не самый важный показатель, поэтому предпочтение отдается дальности передачи. Многие производители умных счетчиков используют для связи проприетарные технологии передачи [2].

II. ИЗМЕРЕНИЕ ПОКАЗАТЕЛЕЙ ЭЛЕКТРОЭНЕРГИИ

Вычисление потребления электроэнергии происходит с помощью измерителя CS5490, который представляет собой аналого-цифровой преобразователь энергии и напряжения. Измеритель вычисляет активную, реактивную, суммарную энергию, реактивные и средние значения напряжения и тока [3]. Связь с измерителем для конфигурации и считывания значений происходит по UART. Стоит отметить, что измеритель является однофазным. Это значит значит, что для измерения трехфазного тока необходимо подключать 3 однофазных измерителя CS5490.

С измерителя CS5490 с настраиваемой периодичностью могут быть считаны следующие данные:

- Энергия активная потребленная
- Энергия активная сгенерированная
- Энергия реактивная потребленная
- Энергия реактивная сгенерированная
- Сила тока в цепи
- Напряжение в цепи
- Фазовый угол
- Частота
- Активная мощность
- Реактивная мощность
- Полная мощность

III. ОПИСАНИЕ ОСНОВНЫХ ПРИНЦИПОВ ВЗАИМОДЕЙСТВИЯ И УПРАВЛЕНИЯ

На Рисунке 2 представлена распиновка измерителя CS5490 для подключения к внешнему управляющему микроконтроллеру

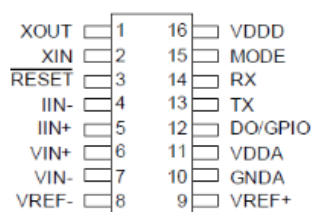


Рис. 2 – Распиновка измерителя CS5490

Дадим описание основных принципов взаимодействия и управления измерителя.

Reset

При установке пина Reset в 1 на время большее, чем 120 нс, происходит прерывание всех операций внутри CS5490 и перезагружается внутреннее состояние и регистры измерителя. После возврата значения пина в 0 начинается инициализационная последовательность, устанавливающая значения по умолчанию для регистров измерителя.

Цифровой выход

Измеритель CS5490 имеет цифровой выход, который может быть настроен на предоставление информации о энергии в виде импульсов,

либо настроен на прерывание по настраиваемым событиям.

Серийный интерфейс UART

CS5490 имеет пин приема и пин отправки для взаимодействия по UART с управляющим микроконтроллером. Измеритель поддерживает асинхронный полнодуплексный UART-порт с 8-битовой передачей. Скорость передачи настраивается конфигурацией измерителя и имеет максимальное значение 512000 бод.

IV. КАЛИБРОВКА ИЗМЕРИТЕЛЯ И КОМПЕНСАЦИЯ ИЗМЕРЕНИЯ

Измеритель CS5490 имеет встроенный алгоритм калибровки. Все вычисления происходят в самом устройстве и записываются во внутренние регистры [4-5]. Однако вычисление компенсации измерения происходит не в самом измерителе, а в управляющем микроконтроллере. Так как измеритель не имеет встроенной энергонезависимой памяти, то для постоянного хранения величин калибровки и компенсации они должны храниться во внешнем управляющем модуле.

Процесс калибровки представляет собой следующий алгоритм:

1. Перезагрузка измерителя CS5490
2. Подача на аналоговый вход заранее известных величин нагрузки
3. Настройка параметров калибровки по нужным показателям (энергия, сила тока, напряжения)
4. Отправка команды старта калибровки
5. Считывание результатов вычисления калибровочных регистров
6. Вычисление компенсаторных значений для новых величин
7. Сохранение значений в энергонезависимой памяти внешнего микроконтроллера

1. A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. [Electronic resource] / J. S. Lee, Y. W. Su, C. C. Shen – 33rd Annual Conference of the IEEE Industrial Electronics Society, 2007. – Mode of access: <https://ieeexplore.ieee.org/document/4460126/> – Date of access: 25.11.2017.
2. Wireless Based Load Control and Power Monitoring System [Electronic resource] / R. Makwana, J. Baviskar, N. Panchal. – International Conference on Energy Efficient Technologies for Sustainability, 2013. – Mode of access: <https://ieeexplore.ieee.org/document/6533559/> – Date of access: 13.07.2018.
3. Two Channel Energy Measurement IC. [Electronic resource] / CIRRUS LOGIC – Datasheet, 2013. – Mode of access: <https://www.cirrus.com/products/cs5490/> – Date of access: 21.01.2018.
4. CS5480/84/90 Measurement Accuracy. [Electronic resource] / CIRRUS LOGIC – Application Note, 2013. – Mode of access: <https://statics.cirrus.com/> – Date of access: 15.03.2018.
5. Remote Power Meter [Electronic resource] / D. Zinner – University of Stellenbosch, 2013. – Mode of access: <http://mtn.sun.ac.za/> – Date of access: 05.10.2017.

ОСОБЕННОСТИ РЕАЛИЗАЦИИ СИММЕТРИЧНЫХ ПУТЕЙ ФНФ ТИПА АРБИТР НА ПЛИС

Иванюк А. А.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ivaniuk@bsuir.by

Предложена новая архитектура звеньев симметричных путей для физической неклонированной функции типа арбитр, позволяющая эффективно использовать аппаратные ресурсы программируемых логических интегральных схем. Показана практическая возможность реализации звеньев симметричных путей на внутренних ресурсах LUT-блоков на примере кристалла FPGA Xilinx Artix-7.

ВВЕДЕНИЕ

Физически неклонированные функции (ФНФ) являются объектами пристального внимания как со стороны научного сообщества, так и со стороны компаний-производителей цифровых устройств и систем, в основном благодаря своим основным свойствам *неклонированности* и *случайности*. Формально произвольную ФНФ можно описать как функцию $PUF(C) = R$, где C определяет множество запросов, а R – множество ответов [1]. Для конкретной реализации ФНФ уникальным и случайным является множество пар запрос-ответ $CR = \{C, R\}$. При этом конкретные значения CR являются случайными и *неуправляемыми*, и становятся известными только после изготовления цифрового устройства либо системы. Реализация идентичной системы (*клонирование*) приведет к получению нового множества $CR^* \neq CR$. Для усиления свойств неклонированности и случайности множество пар CR должно быть достаточно большим и выбирается исходя из практической недостижимости получения значений всех возможных пар. Например, для ФНФ с $N = 128$ входами двоичных значений запросов C мощность множества CR составляет 2^{128} , что делает практически недостижимым их получение, хранение и последующий анализ. Подобные ФНФ называются *сильными* ФНФ, которым приписывают еще одно важное свойство *непредсказуемости*, заключающееся в невозможности определения (прогнозирования) значения конкретной пары $\{C_i, R_i\}$ при известном значении $\{C_j, R_j\}$, $i \neq j$, $i, j \in [0, 2^N - 1]$.

Описанные выше свойства ФНФ применяются для реализации неклонированной идентификации/аутентификации интегральных схем и реализованных на них цифровых устройств, для построения генераторов случайных невоспроизводимых числовых последовательностей, для реализации аппаратных хэш-функций и др. [1].

Однако практическое применение ФНФ сталкивается с рядом проблем, среди которых основной является проблема повторяемости пар CR во времени. Данное свойство определяется такой характеристикой ФНФ как *стабильность*,

значение которой зависит от таких факторов как температура окружающей среды, значение питающего напряжение, степень износа кристалла интегральной схемы и др. Кроме этого такие характеристики ФНФ как уникальность и случайность также являются не идеальными, что заставляет исследователей и разработчиков искать новые методы их улучшения [1].

В данной работе предлагается новая архитектура сильной ФНФ типа арбитр, позволяющая более эффективно использовать имеющиеся конфигурируемые ресурсы ПЛИС. В отличие от существующих архитектур предлагаемая архитектура характеризуется наличием множественных локальных симметричных путей, принципиально неуправляемых со стороны разработчика.

I. ФНФ ТИПА АРБИТР

Схемотехническая реализация ФНФ типа арбитр представляет собой цифровое устройство, включающее в себя следующие основные компоненты: генератор тестового сигнала (ГТС), блок симметричных путей (БСП), состоящий из N симметричных звеньев, и собственно арбитра. Каждое звено имеет два входа для тестового сигнала, один управляющий вход и два выхода, последовательно соединенные с предыдущим и последующим звеном. Таким образом последовательно соединенные N звеньев образуют БСП. Управляющий вход каждого i -го звена соединен с соответствующим входом запроса C_i , значение которого определяет одну из двух возможных конфигураций: $C_i = 0$ – прямая передача значений сигналов со входов звена на соответствующие два выхода и $C_i = 1$ – перекрестная передача значений сигналов. Для N последовательно соединенных звеньев количество возможных комбинаций оценивается как 2^N , что характеризует ФНФ типа арбитр как сильную ФНФ. Два выхода последнего звена, управляемого сигналом запроса C_{N-1} , поступают на схему арбитра, который оценивает какая из копий тестового сигнала от ГТС пришла первой на заданной конфигурации БСП. ФНФ типа арбитр хорошо исследована и апробирована на различных технологиях, в том числе и на кристаллах ПЛИС. Она

обладает достаточной стойкостью к температурным воздействиям, что делает ее пригодной к построению стабильных неклонируемых идентификаторов [1].

При реализации ФНФ типа арбитр на ПЛИС основная доля аппаратных затрат приходится на имплементацию БСП. В общем случае для реализации одного звена необходимо два двухвходовых мультиплексора с перекрестным соединением входных портов. Технологическая схема реализации одного звена на ПЛИС требует наличия двух аппаратных LUT-блоков, а именно LUT-3, для которых использованы три логических входа: два для тестового сигнала и один для сигнала запроса. Обобщая вышесказанное можно утверждать, что для реализации БСП длины N на ПЛИС необходимо использование $2N$ LUT-3 технологических блоков.

Наряду с описанными достоинствами ФНФ типа арбитр обладает и рядом недостатков, среди которых можно отменить имеющуюся линейную зависимость множеств запросов и формируемых на них ответов, позволяющую достаточно эффективно применять методы машинного обучения с целью построения точной программной модели такой ФНФ. Эффективными методами, затрудняющими атаки на ФНФ при помощи машинного обучения, являются архитектурные методы, увеличивающие множество симметричных путей и нелинейные преобразования над множеством запросов и ответов. Это, в свою очередь, приводит к увеличению аппаратных затрат с неэффективным использованием ресурсов ПЛИС.

Современные ПЛИС типа FPGA имеют в своем составе LUT-блоки, позволяющие реализовывать произвольные комбинационные схемы с большим числом входов. Например, FPGA фирмы Xilinx семейства Artix-7 имеет LUT-блоки с шестью входами [2]. В случае реализации одного звена БСП на Artix-7 только три входа из шести LUT-6 блока будут использованы, и будет задействована только 1/8 памяти конфигурации этого же блока.

II. ПРЕДЛАГАЕМАЯ АРХИТЕКТУРА ФНФ ТИПА АРБИТР

В данной работе предлагается новая архитектура симметричных путей ФНФ типа арбитр с учетом их реализации с использованием технологических примитивов LUT-6. Для k входов копий тестового сигнала необходимо наличие $\lceil \log_2 k \rceil$ адресных входов для реализации мультиплексирования сигнала с выбранного входа на единственный выход. Таким образом для LUT-6 блока имеем следующее неравенство $k + \lceil \log_2 k \rceil \leq 6$. Одним из решений данного неравенства является значение $k = 4$. Таким обра-

зом, реализация одного полноценного звена с использованием двух блоков LUT-6 позволяет реализовать четыре конфигурации симметричных путей. Данный подход, в отличие от классической архитектуры, позволяет полностью использовать аппаратные ресурсы LUT-блоков как по числу входов, так и по памяти конфигурации. Если реализация классической схемы ФНФ типа арбитр на FPGA семейства Artix-7 использует $2N$ LUT-6, то предлагаемая архитектура позволяет реализовать ФНФ с большим числом входов при использовании той же аппаратуры.

III. АНАЛИЗ ОСНОВНЫХ ХАРАКТЕРИСТИК

Внутренняя структура блока LUT-6 включает в себя 64-разрядный регистр конфигурации и 63 двухвходовых мультиплексора, формирующие древовидную 6-уровневую схему выборки. По предложенной архитектуре все возможные четыре конфигурации внутри блока LUT-6 формируют локальные симметричные пути прохождения тестового сигнала до единственного выхода. Согласно основному свойству ФНФ все пути в конкретном блоке LUT-6 имеют различные задержки распространения сигнала. Ниже представлены данные о задержках, полученные путем параметрического моделирования VHDL-модели различных звеньев ФНФ типа арбитр, реализованных на различных LUT-блоках кристалла FPGA Artix-7 XC7A100T (CSG324).

Таблица 1 – Значения задержек распространения сигналов для трех различных блоков LUT-6, (ns)

Запрос	LUT-6(1)	LUT-6(2)	LUT-6(3)
00	0,897	0,847	0,838
01	1,023	0,999	0,979
10	0,858	0,927	0,907
11	1,103	1,071	1,044

IV. ЗАКЛЮЧЕНИЕ

Предложена новая архитектура блока симметричных путей для ФНФ типа арбитр, позволяющая экономично использовать внутренние ресурсы ПЛИС типа FPGA. Благодаря использованию локальной неуправляемой симметрии внутренних путей аппаратных LUT-блоков зависимость множества ответов от множества запросов усложняется, что потенциально затрудняет применение атак на данный тип ФНФ с помощью методов машинного обучения.

V. СПИСОК ЛИТЕРАТУРЫ

1. Secure System Design and Trustable Computing / Springer; editors Ch.-H. Chang, M. Potkonjak. — Switzerland, 2016. — p.537.
2. 7 Series FPGAs Configurable Logic Block, User Guide [Electronic resource] / Xilinx Inc., 2016. – Mode of access: <https://www.xilinx.com/>. – Date of access: 08.10.2018.

КОНВЕЙЕРНЫЙ ПРОЦЕССОР ХЭШ-ФУНКЦИИ SHA-256

Качинский М. В., Станкевич А. В.

Кафедра электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {kachinsky, stankevich}@bsuir.by

Рассматривается аппаратная реализация хэш-функции SHA-256 для встраиваемых систем с высокой производительностью. Предлагаемый специализированный процессор SHA-256 имеет конвейерную архитектуру. Приводятся характеристики реализации на базе кристалла ПЛИС XC7VX485T-2FFG1761 фирмы Xilinx.

Хэш-функция SHA-256 [1] относится к семейству SHA-2. Алгоритм SHA-256 практически используется в разнообразных криптографических приложениях. Для приложений, требующих высокой производительности (например, майнеры криптовалют), необходима аппаратная реализация алгоритма. В докладе рассматривается конвейерная реализация алгоритма SHA-256.

Алгоритм вычисления хэш-функции обеспечивает сжатие входного сообщения (для сообщений длиной менее 256 бит – расширение) и получение хэш-значения фиксированной длины 256 бит. Алгоритм заключается в выполнении 64 однотипных циклов (раундов). На рис. 1 приведена схема вычислений для одного раунда.

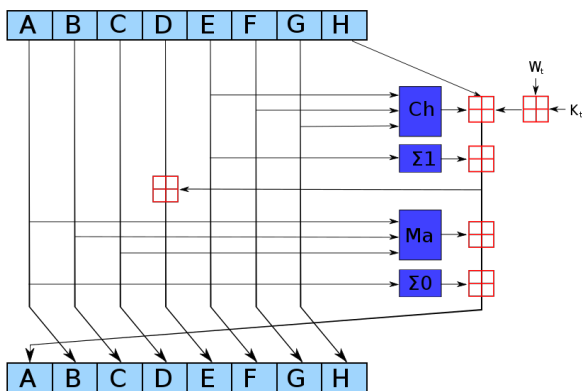


Рис. 1 – Схема вычислений для одного раунда алгоритма SHA-256

На первом раунде 32-разрядные переменные A-H инициализируются константами алгоритма [1]. Входными значениями переменных A-H для следующего раунда являются выходные значения переменных A-H предыдущего, поэтому раунды нельзя вычислять параллельно. Кт является раундовой константой, W_t – раундовым входным словом [1].

В конвейерной архитектуре (рис. 2) блоком данных, продвигающимся по конвейеру является состояние, получаемое на выходе предыдущего процессорного ядра. Каждое процессорное ядро выполняет один раунд алгоритма SHA-256, и эти ядра включены последовательно. Вектор иници-

ализации (IV) подается на вход первого процессорного ядра. Для реализации 64 раундов алгоритма используется цепочка из 64 последовательно включенных процессорных ядер. Параллельно с продвижением текущего состояния по конвейеру продвигается входной блок и текущие данные экспандера для формирования требуемого значения W_t для соответствующего процессорного ядра.

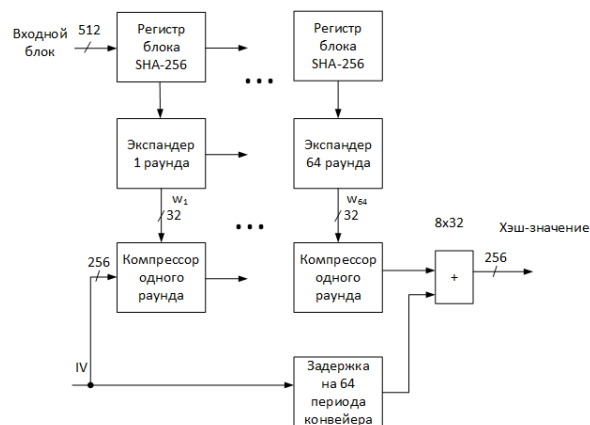


Рис. 2 – Конвейерная архитектура специализированного процессора SHA-256

Выходные данные 64-го процессорного ядра складываются с задержанным на 64 периода (цикла) работы конвейера значением IV для получения выходного хэш-значения.

Если на конвейерный процессор каждый цикл работы конвейера подавать новый входной блок и новый вектор инициализации (в случае, если сообщение занимает более одного блока SHA-256), то по истечении 64 циклов (латентность конвейера) на выходе в каждом цикле работы конвейера будет появляться новое хэш-значение. Если цикл конвейера будет равен одному такту работы вычислительной системы, то производительность конвейерного процессора будет численно равна тактовой частоте. Значение тактовой частоты будет соответствовать числу вычисляемых хэш-значений в секунду.

При конвейерной реализации вычислений одного раунда необходимо принять меры для уменьшения непрерывной цепочки сложений (рис. 1), которая будет определять критический

путь. Для этого следует учесть некоторые особенности алгоритма SHA-256: значения переменных B,C,D после выполнения текущего раунда совпадают со значениями переменных A,B,C до выполнения раунда, а значения переменных F,G,H – совпадают со значениями переменных E,F,G. Эту особенность можно проследить и на несколько раундов назад. Такая особенность алгоритма позволяет выполнить часть вычислений, использующих указанные переменные не в текущем, а в предыдущем такте, либо еще на такт раньше. Также можно формировать значение Wt не в текущем такте, а на такт или два раньше для синхронизации с вычислением новых значений переменных A и E.

Еще одна особенность алгоритма связана с возможностью разнесения на разные такты вычисления переменных E и A, поскольку при вычислении A частично используются промежуточные вычисления для E.

Конвейерная реализация выполнена с учетом указанных особенностей.

На первой ступени конвейера компрессора организуются предварительные вычисления сумм $Wt+Kt+H$ и $Wt+Kt+H+D$, вторая ступень дополнительно вычисляет значение E и значение переменной $T1 = H + BSIG1(E) + CH(E,F,G) + Kt + Wt$ [1] для вычисления A на следующей ступени. Третья ступень является первой полной ступенью конвейера и она дополнительно к указанным выше действиям вычисляет значение A. Таким образом на полной ступени конвейера происходит вычисление значения переменной A текущего раунда и значения переменной E для следующего раунда.

На предпоследней ступени конвейера вычисляется значение A для последнего 64 раунда алгоритма. На последней ступени происходит сложение со значением вектора инициализации алгоритма. С учетом предварительных вычислений и завершающего сложения полный конвейер имеет 67 ступеней и спустя 67 тактов на его выходе появится первое хэш-значение. Далее конвейер будет каждый такт формировать новое хэш-значение при соответствующих изменениях входных данных.

Для разных ступеней конвейера компрессора с помощью конвейера экспандера формируются значения Wt. Для первых 16 раундов (раунды 1 - 16) экспандер реализует задержку входного блока SHA-256 и в качестве Wt выдает соответствующее 32-разрядное слово входного блока. Одна ступень конвейера экспандера для раундов 17-56 дополнительно содержит средства для расширения входного сообщения и вычисления Wt в соответствии с [1]. При реализации конвейера экспандера для раундов с 57 по 64 учтено то обстоятельство, что не все элементы входного бло-

ка данных будут использованы при вычислениях текущего значения Wt.

Аппаратные затраты конвейерного процессора SHA-256 после процедуры синтеза средствами ISE 14.7 для кристалла FPGA XC7VX485T-2FFG1761 приведены на рис. 3. По отчету средств синтеза максимальная тактовая частота составляет 323 МГц.

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	27478	607200	4%
Number of Slice LUTs	35160	303600	11%
Number of fully used LUT-FF pairs	15936	46702	34%
Number of bonded IOBs	1028	700	146%
Number of BUFGBUFCTRLs	1	32	3%

Рис. 3 – Аппаратные затраты предлагаемого конвейерного процессора SHA-256 после процедуры синтеза для кристалла FPGA XC7VX485T-2FFG1761

Для сравнения с предлагаемым процессором была выполнена реализация конвейерного ядра SHA-256 свободно распространяемого проекта Bitcoin-майнера [2] на таком же кристалле FPGA XC7VX485T-2FFG1761. Аппаратные затраты этого конвейерного процессора после процедуры синтеза приведены на рис. 4. По отчету средств синтеза Xilinx ISE 14.7 максимальная тактовая частота составляет 178 МГц.

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	88419	607200	14%
Number of Slice LUTs	61058	303600	20%
Number of fully used LUT-FF pairs	12256	137221	8%
Number of bonded IOBs	3	700	0%
Number of BUFGBUFCTRLs	1	32	3%
Number of DCM_ADVs	1	0	

Рис. 4 – Аппаратные затраты конвейерного процессора SHA-256 [2] после процедуры синтеза для кристалла FPGA XC7VX485T-2FFG1761

По сравнению с свободно распространяемым проектом Bitcoin-майнера [2] предлагаемая реализация SHA-256 почти в два раза производительнее и требует примерно в два раза меньше аппаратных ресурсов кристалла FPGA.

Применительно к процессу майнинга криптовалют можно провести дополнительную минимизацию аппаратных затрат процессора за счет учета структуры заголовков блоков криптовалют.

1. National Institute of Standards and Technology (NIST), FIPS Publication 180-3: Secure Hash Standard, October 2008 [Электронный ресурс] – Режим доступа: <https://csrc.nist.gov/publications/detail/fips/180/3/archive/2008-10-31>. – Дата доступа: 18.09.2018.
2. Bitcoin-miner – [Электронный ресурс]. – Режим доступа: <https://github.com/fpgaminer/Open-Source-FPGA-Bitcoin-Miner/tree/master/projectst>. – Дата доступа: 18.09.2018.

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ АЛГОРИТМОВ ДИЗЪЮНКТИВНО-КОНЪЮНКТИВНОГО РАЗЛОЖЕНИЯ СИСТЕМ ПОЛНОСТЬЮ ОПРЕДЕЛЕННЫХ БУЛЕВЫХ ФУНКЦИЙ

Кардаш С. Н.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: gold@newman.bas-net.by

Приводятся результаты экспериментального исследования эффекта от влияния применения процедур, основанных на построении дизъюнктивно-конъюнктивных разложений систем полностью определенных булевых функций, на площадь получаемых в результате синтеза схем из библиотечных КМОП-элементов. Показано, что сокращение площади схем в результате применения предлагаемых программ может достигать 25 процентов

ВВЕДЕНИЕ

Синтез нерегулярных логических схем из библиотечных КМОП элементов представляет собой двухэтапный процесс [1]. Первый этап заключается в логической оптимизации – исходная система булевых функций записывается в виде взаимосвязанных логических уравнений, которые называют многоуровневым представлением булевых функций, в отличие от двухуровневых (И-ИЛИ) представлений в виде ДНФ. Оптимизация многоуровневых представлений может производиться различными способами. В настоящей работе она осуществляется на основе поиска минимизированных BDD-представлений (Binary Decision Diagram) систем булевых функций [2]. Цель этапа оптимизации заключается в построении такого варианта представления схемы, который мог бы послужить хорошей отправной точкой для синтеза. После этапа оптимизации выполняется второй этап – технологическое отображение оптимизированного многоуровневого BDD-представления в заданный базис библиотечных логических элементов, выполненных по КМОП-технологии. Для этого используется программный пакет LeonardoSpectrum [3]. В результате получается структурное VHDL-описание (нетлист) комбинационной схемы, реализующей исходную систему функций. Основными критериями при оптимизации схем являются число ячеек (cell) и площадь кристалла, занимаемого полученным блоком (square).

1. ОРГАНИЗАЦИЯ ЭКСПЕРИМЕНТА

В работе [4] рассмотрен алгоритм построения совместных дизъюнктивных и конъюнктивных разложений систем булевых функций, ориентированный предварительную обработку исходных систем функций, обладающих связностью (общностью) областей определения. Основная идея алгоритма заключалась в том, чтобы находить и выделять общие части логических выражений с учетом специфики используемой целевой библиотеки (реализовать можно

как функции, так их инверсии). В этом случае процесс синтеза можно представить как трехэтапный процесс – построение дизъюнктивно-конъюнктивного разложения, минимизация в классе BDD с использованием программ [2] и отображение в библиотечный базис средствами программного пакета LeonardoSpectrum. В настоящей работе приводятся результаты экспериментального исследования эффекта от применения двух программ, разработанных на основе алгоритма [4]. Программы различаются способом учета полярности включаемых в разложение функций. Программа A1 строит разложение без учета полярности, а программа A2 полярность учитывает.

Исследование проводилось в системе FLC [5]. В качестве объектов исследования использовались функциональные описания (системы полностью определенных булевых функций) проектируемых схем, представленные на языке SF [6], который является внутренним языком системы FLC. Исходные функциональные описания схем взяты из библиотеки примеров Berkeley PLA Test Set [7]. Для каждого примера применялась многошаговая процедура. На первом шаге проводилась построение дизъюнктивно-конъюнктивного разложения. В результате получалось многоблочное представление на языке SF. На втором шаге оно минимизировалось в классе BDD с привлечением механизма стратегий [5]. В результате получалось многоуровневое SF-описание, каждый уровень которого задавался скобочным выражением. После этого осуществлялась конвертация многоуровневых SF-описаний в синтезируемые алгоритмические VHDL-описания. Получение схемного решения и его оценка выполнялась с помощью программного пакета LeonardoSpectrum. Всего для каждого примера рассматривалось три варианта логической оптимизации. При первом – программы разложения не использовались. Во втором и третьем случае вначале применялись программы A1 и A2 соответственно.

II. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ

Результаты экспериментального исследования представлены в таблице 1. Всего для каждого примера рассматривалось три варианта логической оптимизации. При первом – программы построения разложения не использовались. Этому варианту в таблице соответствует столбцы с обозначением А. Во втором и третьем случае вначале применялись программы А1 и А2 соответственно. В приведенной таблице через m обозначено число аргументов в исходной системе, через cells - число макроячеек в полученной схеме, через square - площадь схемы. Как следует из приведенной таблицы, использование разложений принесло эффект в десяти случаях из 15. При этом наибольшее сокращение площади составило 26 процентов (в случае примера intb).

III. ЗАКЛЮЧЕНИЕ

Уменьшить площадь получаемых в процессе синтеза комбинационных нерегулярных логических схем можно с помощью логической оптимизации исходных описаний. Одним из наиболее эффективных способов логической оптимизации является минимизация BDD-представлений систем булевых функций. Хорошим средством для улучшения конечных схемных решений может служить предварительная обработка исходных систем с помощью программ построения дизъюнктивно-конъюнктивных разложений. Проведенное исследование показало, что более, чем

в половине исследованных случаев применение разработанных программ приводило к получению более эффективных схемных решений. При этом использование разработанных программных средств может сократить площадь получаемых схем на 25 процентов.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Бибило, П.Н. Логическое проектирование дискретных устройств / П.Н. Бибило, В.И. Романов. – Минск.: Беларус. навука, 2011. – 279 с.
2. Бибило П.Н., Леончик П.В. Экспериментальное исследование влияния процедуры BDD-оптимизации на энергопотребление комбинационных КМОП-схем // Автоматика и вычислительная техника. 2010. № 5. С. 72-78.
3. Бибило П.Н. Системы проектирования интегральных схем на основе языка VHDL. StateCAD, ModelSim, LeonardoSpectrum. – М.: СОЛОН-Пресс. 2005, 384 с.
4. Бибило, П.Н. Дизъюнктивно-конъюнктивные разложения систем полностью определенных булевых функций / П.Н. Бибило, С.Н. Кардаш / Доклады Восьмой Международной научной конференции «Танаевские чтения», 27–30 марта 2018 г. – Минск, ОИ-ПИ НАН Беларуси, 2018 г. – С. 28–32.
5. Бибило П.Н., Романов В.И. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний. Изд. 2-е, испр. – М.: ЛЕНАНД, 2014, 256 с.
6. Бибило, П.Н. Кремниевая компиляция заказных СВИС. - Минск: Институт технической кибернетики АН Беларуси, 1996. - 268 с.
7. Berkeley PLA test set [Electronic resource]. – Mode of access: <http://www1.cs.columbia.edu/cs4861/sis/espresso-examples/ex/>. – Date of access: 10.10.2011.

Таблица 1 – Результаты эксперимента

Имя	m	A	A1	A2	A	A1	A2
		cells	cells	cells	square	square	square
Add6	12	28	42	59	12806	18425	24524
Addm4	8	212	187	224	80782	73723	88426
B2	16	517	830	830	192655	319438	319438
B12	15	50	52	47	18297	18911	16645
DC2	8	65	58	58	23302	21455	21455
intb	15	711	650	534	272555	261691	215734
Mp2d	14	52	52	52	17471	18018	18018
radd	8	18	20	17	8074	9151	8576
root	5	72	68	65	26109	24396	24625
T3	8	49	47	47	17276	16941	16941
tial	14	644	770	879	255988	295885	355742
Z5xp1	7	48	44	50	18442	17209	18833
Max10	10	390	345	402	146888	130823	156084
Max5	9	221	211	197	84643	78025	76904
F51	8	44	43	43	18353	17599	17599

ПЕРСПЕКТИВНЫЕ ВОЗМОЖНОСТИ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ИНФРАСТРУКТУРЫ IOT

Клыбик В. П., Иванюк А. А.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: vold029@gmail.com, ivaniuk@bsuir.by

Рассматриваются перспективы применения физически неклонлируемых функций для идентификации и аутентификации устройств интернета вещей (IoT).

ВВЕДЕНИЕ

Интернет вещей (ИВ) (Internet of Things - IoT) — концепция, подразумевающая взаимодействие посредством сетевых коммуникаций с помощью стандартных интернет-протоколов между кибер-физическими системами, являющимися разнообразными вещами физического мира, например приборы учета электроэнергии, датчики температуры, влажности и др.

Подобные системы существовали и ранее, но ограничивались географической распределенностью, высокой стоимостью, ограниченностью коммуникаций и зонами использования. Прогресс в микроэлектронике и развитие коммуникаций привели к бурному росту количества устройств ИВ и их глубокое проникновение в различные сферы. Gartner заявляет, что в 2017 году число подключенных устройств составило 8,4 млрд [1]. Такое быстрое и широкое распространение ИВ с одной стороны дает множество новых возможностей для развития каждой сферы, с другой стороны - обостряет ряд проблем, без решения которых дальнейшее проникновение ИВ может привести к катастрофическим последствиям.

I. ИНФРАСТРУКТУРА ИНТЕРНЕТА ВЕЩЕЙ

В общем случае ИВ является лишь концепцией взаимодействия посредством сетевых коммуникаций между разнообразными устройствами. Следствием для всей отрасли является высокая гетерогенность решений, отсутствие жесткой структуры, архитектуры и границ в критериях классификации.

На формирование практической инфраструктуры влияют следующие факторы:

- многие устройства ИВ крайне ограничены в вычислительных ресурсах;
- устройства общаются не только, а, иногда, и не столько между собой, сколько со специальными приложениями/сервисами, реализующими высокоуровневую сервисную логику управления или обработки полученных данных;
- прямое подключение к сети Интернет невозможно, т.к. не позволяет реализовать требования по скорости отклика/обмена,

или необходимо использовать несовместимые протоколы связи.

Исходя из изложенного выше, развернутая структура ИВ представлена на Рис.1. Стрелки показывают направления управляющих и информационных потоков данных между взаимодействующими компонентами системы.

Например, компания Cisco предложила семиуровневую архитектуру ИВ[2].

II. ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ ИНФРАСТРУКТУРЫ

В любом варианте инфраструктуры совершенно важными факторами являются идентификация и аутентификация каждого компонента. Идентификация необходима в первую очередь для корректной адресации потоков управления и ассоциации данных. Аутентификация позволяет убедиться, что не произошла несанкционированная подмена компонента.

Например, в случае умного дома может быть по ошибке или злонамеренно изменена температура помещений, включена вода, свет, бытовые приборы. Как минимум, это вызовет сильный дискомфорт у человека. В случае промышленного применения ИВ последствия могут быть серьезными.

На сегодняшний день используются статический и динамический способ идентификации компонентов. Статический способ реализуется с использованием:

- прошитых идентификаторов;
- MAC-адресов;
- QR-кодов.

Динамический способ использует наблюдение и анализ сетевого трафика компонента.

В случае статического способа идентификации одной из проблем является обеспечение уникальности идентификаторов и возможность дублирования идентификаторов устройств в случае ошибки при производстве или настройке.

Динамический способ идентификации требует длительного обучения анализатора, имеет низкую стабильность и практически не применим на этапе первичного ввода компонента в эксплуатацию.

Для аутентификации устройств используются крипто-функции и крипто-протоколы, основанные на устанавливаемых на компоненты

сертификатах. Сертификат представляет собой специальный набор данных, хранимый во внутренней памяти устройства. Для успешного использования крипто-протоколов при аутентификации важными являются две проблемы:

- обеспечить неизвлекаемость сертификата из внутренней памяти устройства;
- обеспечить генерацию истинно случайных последовательностей чисел.

На сегодняшний день оба требования либо недостаточно обеспечиваются в случае массовых дешевых устройств ИВ, либо требуют применения дополнительных специализированных аппаратных решений, что ведет к удорожанию устройств [3,4].

III. ПРИМЕНЕНИЕ ФНФ

Формальное определение физически неклонированной функции (ФНФ) цифрового устройства дано в работе [5]. Основными свойствами ФНФ являются:

- невозпроизводимость математической/алгоритмической модели;
- не копируемость при тиражировании схемной реализации.

Основными применениями ФНФ являются:

- идентификация цифровых устройств;
- генерирование крипто-ключей.

Достоинствами применениями ФНФ для ИВ является отсутствие:

- необходимости в специализированных аппаратных решениях;
- процесса назначения идентификатора устройства при производстве.

В текущее время большинство практических реализаций ФНФ основаны на программируемых логических интегральных схемах (ПЛИС), либо в составе специализированных процессоров. В случае ИВ часто используются широко распространенные микроконтроллеры. Создание реализаций ФНФ для микроконтроллеров является

важной задачей для применения в ИВ. Такие ФНФ позволят повысить безопасность не только новых, но и существующих устройств ИВ, путем обновления программного обеспечения.

Выводы

С учетом роста проникновения ИВ в многие сферы жизнедеятельности человека, обеспечение безопасности становится критическим фактором успешного развития отрасли.

Применение ФНФ в задачах идентификации и аутентификации ИВ является перспективным направлением для дальнейших исследований. Необходимо как исследовать реализации ФНФ на текущем аппаратном обеспечении ИВ на предмет достаточного качества их характеристик, так и предложить реализации ФНФ в будущих версиях аппаратного обеспечения.

ЛИТЕРАТУРА

1. Gartner portal / Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016 // [Electronic resource]. – Mode of access: <https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says-8-billion-connected-things-will-be-in-use-in-2017-up-31-percent-from-2016>. – Date of access: 01.10.2018.
2. Internet of Things Word Forum / Internet of Things Word Forum [Electronic resource]. – Mode of access: http://cdn.iotwf.com/resources/71/IoT_Reference_Model_White_Paper_June_4_2014.pdf. – Date of access: 02.10.2018.
3. State-of-the-art answers to today's embedded security challenges with OPTIGA Trust / infineon Technology portal [Electronic resource]. – Mode of access: <https://www.infineon.com/cms/en/product/security-smart-card-solutions/optiga-embedded-security-solutions/optiga-trust/>. – Date of access: 02.10.2018.
4. iBadge / InfoKeyVault portal [Electronic resource]. – Mode of access: <http://www.ikv-tech.com/index.php/en/product/ibadge>. – Date of access: 02.10.2018.
5. Ярмолик, В. Н. Физически неклонированные функции / В. Н. Ярмолик, Ю. Г. Вашилко // Информатика. – 2011. – №2. – С. 90-100.

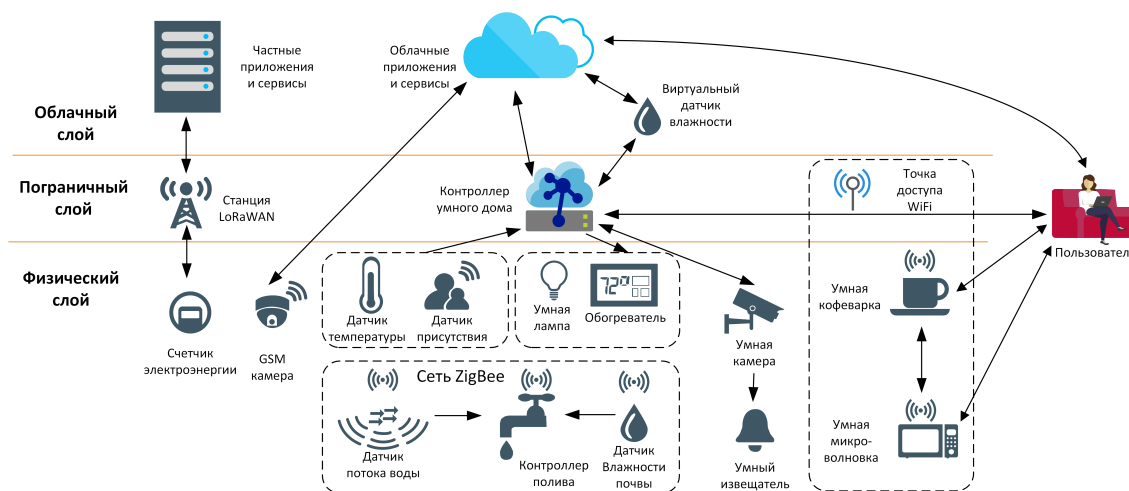


Рис. 1 – Развернутая структура системы интернета вещей

ИСПОЛЬЗОВАНИЕ БИБЛИОТЕК LIBERTY ПРИ ПОСТРОЕНИИ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ДЛЯ ОЦЕНКИ ЭНЕРГОПОТРЕБЛЕНИЯ КМОП-МИКРОСХЕМ

Королёв Н. А.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: kolkor1@yandex.ru

Рассматриваются основные положения формата LIBERTY. Приводится пример построения математической модели для расчета потребления энергии КМОП-микросхем с помощью справочных таблиц формата.

ВВЕДЕНИЕ

Формат LIBERTY – стандартизированный промышленный формат, который используется для создания библиотек, в которых описываются различные элементы, объединенные одной технологией [1]. Описание элементов включает в себя различную информацию: временных характеристики, мощностные характеристики, площадь, функции, рабочие диапазоны и т. д. Формат LIBERTY является текстовым документом, который чаще всего представляется в виде текстового ASCII-файла с расширением «.lib». Элементы, которые описываются в файлах библиотеки, моделируются при различных условиях, отражающих реальные условия, в которых может оказаться элемент (ячейка). Подобный процесс называют характеристикой элементов библиотеки. Например, задержка на выходе инвертора зависит от длины входного фронта и суммарной емкостной нагрузки на выходе. Инвертор будет характеризоваться набором входных фронтов и выходных емкостных нагрузок. Полученные в результате характеристики параметры будут представлены в виде справочных таблиц, которые будут содержать в себе значения времени задержек инвертора в зависимости от различных протяженностей входных фронтов и величин емкостных нагрузок. Формат LIBERTY включает в себя различные методы характеристик моделей элементов: модель нелинейной задержки (NLDM), модель композитного источника тока (CCS), модель источника эффективного тока (ECSM).

I. Мощность

Характеристики мощности – важная составляющая характеризующих моделей. Рассеиваемая мощность состоит из трех компонентов: мощности утечки, внутренней мощности и мощности переключений. Мощность утечки. Мощность, которая рассеивается ячейкой, когда она не активна. Эта утечка возникает за счет небольших пороговых напряжений, которые не позволяют ячейке окончательно выключиться. Внутренняя мощность. Во время переключения цепь рассеивает внутреннюю мощность, заряжая или

разряжая любую емкость, существующую внутри ячейки, однако изменения напряжения не всегда приводят к логическому переключению. Справочные таблицы, которые содержат значения внутренней мощности, имеют названия `rise_power` и `fall_power`. Мощность переключений или емкостная мощность – это мощность, рассеиваемая при зарядке и разрядке емкостной нагрузки на выходе ячейки. Переключательная активность – величина, которой измеряют количество переключений (из 0 в 1 и из 1 в 0) для каждого соединения между элементами цепи, когда приходит входной импульс. Для удобства моделирования переключательная мощность рассматривается как компонент внутренней мощности.

II. ВРЕМЕННЫЕ ХАРАКТЕРИСТИКИ

Временные характеристики (рис. 1) являются неотъемлемой частью модели нелинейных задержек, как и самого формата LIBERTY. Для моделей нелинейных задержек характерны следующие временные характеристики: задержка распространения, задержка ячейки и время переключения. Время переключения. Справочные таблицы, которые содержат в себе значения времени переключения для NLDM моделей, называются `rise_transition` и `fall_transition`. Время переключения (transition time) – время, которое необходимо для переключения из 0 в 1 (`rise_transition`) или из 1 в 0 (`fall_transition`). Задержка переключения является функцией суммарной емкости на выходе и времени переключения на входе. Задержка распространения. Справочные таблицы, которые содержат в себе значения времени задержки распространения для NLDM моделей, называются `rise_propagation` и `fall_propagation`. Задержка распространения – задержка между порогами переключения (верхний порог для переключения из 0 в 1 и нижний – для переключения из 1 в 0) на входе и выходе ячейки, она включает в себя время переключения. Задержка распространения является функцией суммарной емкости на выходе и времени переключения на входе элемента. Задержка ячейки. Справочные таблицы, которые содержат в себе значения времени задержек ячеек для NLDM

моделей, называются cell_rise и cell_fall. Важно отметить, что при наличии таблиц задержек ячеек не используются таблицы задержек распространения (propagation_rise, propagation_fall), и наоборот. Задержки ячеек определяются независимо от времени переключения. Задержка ячейки является функцией суммарной ёмкости на выходе и времени переключения на входе.

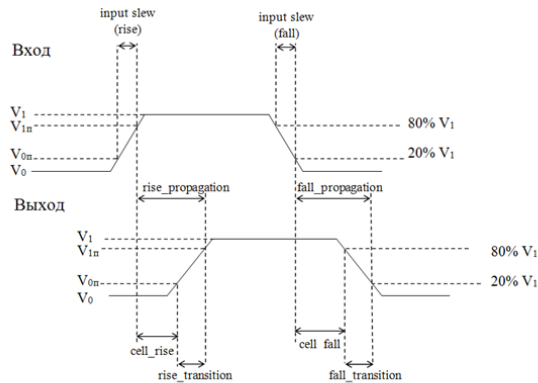


Рис. 1 – Временные характеристики логических элементов

III. ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ С ПОМОЩЬЮ СПРАВОЧНЫХ ТАБЛИЦ LIBERTY

Рассмотрим логическую схему (рис. 2), состоящую из шести логических элементов [2]. Чтобы определить мощность любого логического элемента при переключении входных либо внутренних сигналов необходимо знать межсоединения всех входов и выходов. Это позволит определить значения ёмкостей на выходах, а также пути распространения сигналов. Искомые значения вычисляются с помощью справочных таблиц формата Liberty.

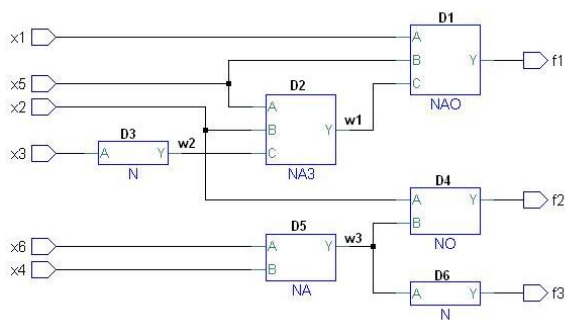


Рис. 2 – Логическая схема из шести логических элементов

Для вычисления рассеиваемой мощности при переключении из 0 в 1 элемента D3 (инвертор) нам необходимо знать величину длительности входного фронта сигнала X3, а также суммарную ёмкость на выходе элемента. Поскольку сигнал X3 приходит извне, то его дли-

тельность мы задаем как параметр моделирования. Величина суммарной ёмкости на выходе элемента D3 в данном случае – величина ёмкости входа C элемента D2. Длительность входного фронта примем равной 0,2 нс, $C_{вх} = C_{вх_C_NA3} = 0.0176298$ пФ. Таблица из формата Liberty (рис.3), содержит значения величины мощности, рассеиваемой элементом НЕ при переключении из 0 в 1.

```

rise_power ( ier_6_x_cl_6_power ) {
-- длительность входного фронта
  index_1("0.045 & 0.09 & 0.18 & 0.35 & 0.5 & 1.4");
-- ёмкостная нагрузка
  index_2("0.0175 & 0.05 & 0.09 & 0.16 & 0.3 & 0.6");
-- рассеиваемая мощность
  values ( \
0.251778 & 0.407236 & 0.594155, 0.918697, 1.57018, 2.9594", \
0.254336 & 0.40565 & 0.593081, 0.919386, 1.56868, 2.95887", \
0.270846 & 0.414878 & 0.596213, 0.920595, 1.56699, 2.96111", \
0.310805 & 0.447798 & 0.622583, 0.938206, 1.57542, 2.95636", \
"0.35487, 0.485616, 0.6383, 0.959954, 1.59515, 2.9719", \
"0.627549, 0.738795, 0.889187, 1.15425, 1.75964, 3.08147" ); }

```

Рис. 3 – Справочная таблица Rise_power

Поскольку в таблице нет значения, которое подходило бы под заданные параметры (0,2 нс, $C_{вх}$), необходимо интерполировать искомое значение по четырем ближайшим точкам. Для этого можно применить метод билинейной интерполяции [3], либо воспользоваться встроенными функциями MathCAD. Ближайшими точками будут значения $Q_{11}=0.270846$ мВт, $Q_{12}=0.310805$ мВт, $Q_{21}=0.414878$ мВт, $Q_{22}=0.447798$ мВт. В результате интерполяции получим значение $Q=0.276119$ мВт. Сигнал X3, проходя через элемент D3, не только инвертируется, но и изменяет длительность фронта. Интерполируем значение длительности фронта сигнала W2 по четырем ближайшим точкам из справочной таблицы для времени переключения. $T_{ft}=0.086675$ нс. Это значение будет использоваться при расчетах для входа C элемента D2. Для построения более точной модели необходимо также рассчитать время переключения внутри ячейки. Оно производится таким же образом. В зависимости от поставленных задач формат LIBERTY позволяет оперировать как временными, так и мощностными характеристиками логических элементов. Существует возможность составлять модели разной степени точности в зависимости от используемых справочных таблиц.

1. Synopsys. Liberty User Guide [Electronic resource] – Mode of access: https://media.c3d2.de/mgoblin_media/media_entries/659/Liberty_User_Guides_and_Reference_Manual_Suite_Version_2017.06.pdf/ – Date of access: 20.09.2018.
2. Бибило, П. Н. Система логического проектирования функциональных блоков заказных КМОП СБИС с пониженным энергопотреблением / П. Н. Бибило, Н. А. Авдеев, С. Н. Кардаш, Н. А. Кириенко, Ю. Ю. Ланкевич, И. П. Логинова, Д. И. Черемисинов, Л. Д. Черемисинова // Микроэлектроника, – 2018, – № 1 – С. 72–88.
3. Cambridge University Press, Numerical Recipes in C. – 1988–92. – P. 123–128.

БИОМЕТРИЧЕСКАЯ ИЗМЕРИТЕЛЬНАЯ СИСТЕМА ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ ПК

Костюк Д. А., Латий О. О., Маркина А. А.
Кафедра электронных вычислительных машин и систем,
Брестский государственный технический университет
Брест, Республика Беларусь
E-mail: dmitriykostiuk@gmail.com

В работе рассматривается микропроцессорная система для биометрического мониторинга и контроля физического состояния пользователя ПК во время его работы с различными программными продуктами и периферийными устройствами. К числу измеряемых параметров относятся сердечный ритм, электрическая активность кожи, кинематическая активность. Рассматривается архитектура устройства и методика измерений.

ВВЕДЕНИЕ

Измерение физиологического состояния пользователя ПК во время работы с программным обеспечением позволяет выполнять экспресс-оценку эффективности человеко-машинного взаимодействия (ЭЧМВ): биометрические данные, полученные в ходе эксперимента, допускают автоматическую интерпретацию, на основе которой в свою очередь может быть выполнена оценка физической и когнитивной нагрузки пользователя. К числу информативных параметров организма можно отнести сердечный ритм, кровяное давление, электропроводность кожи, активность мышц и головного мозга, причем достоверная оценка предполагает одновременное измерение нескольких параметров [1, 2] с отслеживанием их корреляции.

I. ОБЩАЯ АРХИТЕКТУРА СИСТЕМЫ

Разработанная система схематично представлена на рисунке 1.



Рис. 1 – Структура системы

Аппаратные измерительные модули – в зависимости от конфигурации датчиков, это модули частоты сердечных сокращений (ЧСС), кинематической активности (КА) и электрической активности кожи (ЭАК) – выполняют помимо измерения также фильтрацию и усиление электрических сигналов. Микроконтроллер осуществляет сбор показаний с датчиков и передаёт поток данных через USB либо Bluetooth-соединение [3] в ПК, выполняющий первичную обработку, визуализацию и журналирование. Программное обеспечение ПК рассчитано как на автономное использование, так и на работу в составе распределенной системы тестирования ЭЧМВ [2], выполняющей параллельный

сбор, централизованное хранение и оценку биометрических данных в ходе эксперимента.

II. ОСОБЕННОСТИ ИЗМЕРЕНИЙ ЧСС

ЧСС измеряется по принципу рефлекторной фотоплетизмографии (ФПП), за счет отслеживания интервалов изменения оптической плотности тканей с помощью источника света и фотодетектора [3] по следующей формуле:

$$p = 60000/t_{peak}, \quad (1)$$

где 60000 – количество миллисекунд в одной минуте, t_{peak} – время между ударами пульса, мс.

ЧСС определяется по следующему алгоритму. Значение базовой линии B_l приравнивается к текущему значению ФПП-сигнала с добавкой, равной 1/10 максимально возможного значения. Значения опорного уровня Ref , а также минимальное Min и максимальное Max значения ФПП-сигнала в пределах одного удара в момент времени $t = 0$ приравниваются B_l и сбрасывается флаг, соответствующий наличию детектированного удара ($Beat = 0$).

На каждой итерации осуществляется поиск реальных значений Min и Max , при условии выполнения следующего неравенства:

$$t_{min} < t_{prev} < t_{max}, \quad (2)$$

где t_{prev} – время, прошедшее после последнего обнаруженного удара пульса, а t_{min} и t_{max} – физиологически-обусловленные минимальный и максимальный интервалы между пиками пульса.

Если текущий уровень ФПП-сигнала $Cur < Ref$ и $Cur < Min$, а также выполняется условие (2), к Cur приравнивается значение Min , а если $Cur > Ref$ и $Cur > Max$, к нему приравнивается значение Max . Если $Cur > Ref$, выполняется условие (2), а флаг удара пульса не установлен, то детектируется удар ($Beat = 1$), и рассчитывается значение ЧСС по формуле (1).

Далее, если $Cur < Ref$ и $Beat = 1$, рассчитывается размах импульса $A = Max - Min$, опорный уровень $(A/2 + Min)$, а Max и Min приравниваются к Ref . Если время с последнего удара

больше t_{max} , значения Ref , Min и Max приравниваются к значению Bl . На каждой итерации рассчитывается Bl , а также t_{prev} .

III. ИЗМЕРЕНИЕ КА

КА измеряется датчиками линейного ускорения (акселерометры) в составе инерциальной измерительной системы InvenSense MPU6050 [4]. Расчёт перемещений выполняется методом интегрирования ускорения [5]:

$$x_n = x_{n-1} + \frac{(v_{n-1} + v_n) \cdot \Delta t}{2}$$

К сигналу применяется низкочастотный фильтр (ФНЧ), исключающий из показаний акселерометра высокочастотные шумы, а также, для устранения составляющей гравитационного притяжения, высокочастотный фильтр (ФВЧ):

$$g_{xyz}^{lpf} = \alpha \cdot g_{xyz} + (1 - \alpha) \cdot a_{xyz}; a_{xyz}^{hpf} = a_{xyz} - g_{xyz}$$

Коэффициент фильтра α рассчитывается по формуле $\alpha = \frac{t}{t+dT}$, где t – временная константа фильтра, dT – период дискретизации сигнала.

Таким образом, упрощенный алгоритм определения перемещения сводится к следующим шагам: считывание показаний акселерометра; устранение высокочастотных помех и удаление составляющей гравитационного ускорения; расчет скорости интегрированием и устранение неточности интегрирования; расчет перемещения посредством интегрирования.

IV. ИЗМЕРЕНИЕ ЭАК

ЭАК как параметр тесно связана с автономной эмоциональной и когнитивной обработкой и широко используется как показатель для анализа симпатической активности [6].

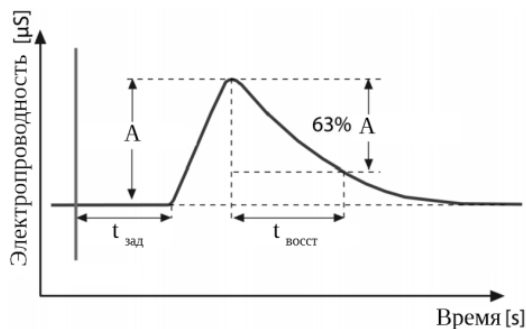


Рис. 2 – Реакция проводимости кожи на раздражитель

Две основные составляющие ЭАК – тоническая и фазическая. Тонический уровень связан с более медленными компонентами и фоновыми характеристиками сигнала (общий уровень, медленное изменение с течением времени). Наиболее распространенной мерой этого компонента является уровень проводимости кожи (УПК), и изменения в УПК, которые отражают общие измене-

ния в вегетативном возбуждении. Фазическая составляющая относится к быстро изменяющимся элементам сигнала, т.е. отражает реакцию проводимости кожи (РПК). При этом фазовый РПК часто получает наибольший интерес, но составляет лишь небольшую часть ЭАК.

Для выделения эмоциональных всплесков из сигнала ЭАК в системе применён следующий алгоритм. Флуктуации сигнала уменьшаются цифровым ФНЧ с частотой среза 2 Гц с конечной импульсной характеристикой [7]:

$$y[n] = b_0x[n] + b_1x[n-1] + \dots + b_Nx[N]$$

где N – порядок фильтра. Далее вычисляется первая производная для выделения пиков сигнала [8], по которым выявляются эмоциональные всплески пользователя и их длительность:

$$y'[n] = y[n] - y[n-1]$$

Для оценки ЭЧМВ применяется нормализованный параметр L – общая длительность эмоциональных всплесков относительно длительности эксперимента t :

$$SCR_t = \sum_{i=0}^N SCR_{ti}; L = SCR_t/t$$

V. СПИСОК ЛИТЕРАТУРЫ

- Schall, A. (2015). The Future of UX Research: Uncovering the True Emotions of Our Users / A. Shall // User Experience Magazine. – 2015. – Vol. 15(2). – Mode of access: <http://uxpamagazine.org/the-future-of-ux-research/>. – Date of access: 30.06.2017.
- Латий, О.О. Подход к биометрической оценке эргономики графического интерфейса пользователя / Д.А. Костюк, О.О. Латий, А.А. Маркина // Вестник БрГТУ: физика, математика, информатика, №5 (101), 2016. – Брест, 2016. – С. 46–49.
- Латий, О.О. Микроконтроллерная система для биометрической оценки состояния пользователя ПК / Д.А. Костюк, О.О. Латий, В.П. Шамолин // Информационные технологии и системы 2017 (ИТС 2017): материалы МНК. Минск, БГУИР, 25 октября 2017 г. – С. 238–239.
- MPU-6000 and MPU-6050 Product Specification Revision 3.4 [Electronic resource] – Mode of access: <https://goo.gl/gmQ36i>. – Date of access: 10.09.2017.
- Латий, О. О. Подход к оценке кинематической активности для анализа эффективности человеко-машинного взаимодействия / О. О. Латий, Д. А. Костюк // Современные проблемы математики и вычислительной техники: сборник материалов X РНК молодых ученых и студентов, Брест, БрГТУ, 23–24 ноября 2017 г. – Брест, 2017. – С. 47–50.
- Martini, F. Essentials of Anatomy & Physiology / F. Martini, E. Bartholomew // San Francisco: Benjamin Cummings, 2003. – P. 267.
- FIR Filter Design by Windowing: Concepts and the Rectangular Window [Electronic resource] – Mode of access: <https://goo.gl/emSJJQz/>. – Date of access: 14.09.2018.
- Differentiation [Electronic resource] – Mode of access: <https://goo.gl/cjPDyS>. – Date of access: 14.09.2018.

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ВИЗУАЛИЗАЦИИ ПОТОКА ДАННЫХ, ПОСТУПАЮЩИХ В ПЭВМ ОТ МИКРОКОНТРОЛЛЕРНОЙ СИСТЕМЫ

Латий О. О.

Кафедра ЭВМ И Систем, Брестский государственный технический университет информатики
Брест, Республика Беларусь
E-mail: latijoo@tut.by

В докладе представлен вариант программного обеспечения для работы с микроконтроллерной системой для биометрической оценки состояния пользователя ПК [1], включающий авторский высокопроизводительный компонент библиотеки Qt для отрисовки 2D-графиков в реальном времени. Приводятся архитектура и особенности реализации, а также результаты сравнения с другими средствами построения графиков для платформы Qt.

ВВЕДЕНИЕ

В ряде задач, связанных с измерительными приборами, способными обмениваться данными с компьютером, возникает необходимость оперативной визуализации получаемых компьютером данных. Жесткие временные ограничения накладывают дополнительные условия на программные компоненты, выполняющие отрисовку графика по поступающим данным. При решении подобной задачи для системы биометрической оценки состояния пользователя [1], автором был разработан кросс-платформенный модуль для библиотеки Qt, выполняющий высококачественную визуализацию таких данных. Код проекта доступен по адресу <https://github.com/lattoo/plotter> под лицензией GPL v.2.

I. АРХИТЕКТУРА СИСТЕМЫ

Структурная схема программного обеспечения, требующего визуализации данных, показана на рисунке 1.

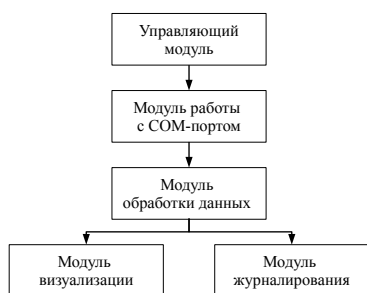


Рис. 1 – Структурная схема программного обеспечения

Эмпирически было получено минимальное значение задержки в 16 мс между принятием данных в режиме работы программного обеспечения в общем потоке. Данная величина является недопустимо большой, поэтому особенностью модуля работы с COM-портом (и его реализацией в виде USB-переходника, через который в настоящее время работают многие источники данных) является запуск подпрограммы получения и распределения данных в отдельном потоке

в целях ускорения взаимодействия между программной частью ПЭВМ и устройством.

На принимающей стороне в зависимости от номера индекса формируются реальные значения, включая выполнение операций сдвига и объединения байт для получения слов. Таким образом удалось достичь контролируемой величины временной задержки в 1 мс.

Модуль обработки данных осуществляет цифровую фильтрацию данных, подсчет средних, минимальных, максимальных величин, отклонений от параметров, а также некоторый статистический анализ полученных результатов.

Модуль журналирования осуществляет запись необходимых данных в csv-файл. Операция записи осуществляется с интервалом 250 мс. Данный временной промежуток является компромиссным, предполагающим получение журналирования результатов с необходимой и достаточной частотой с одной стороны и получение конечного файла относительно небольшого объема, с другой.

Интерфейс можно увидеть на рисунке 2.

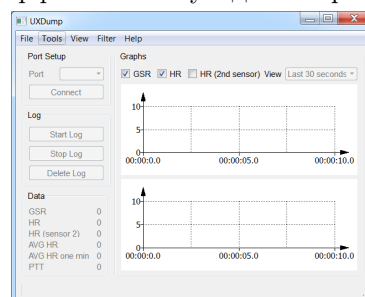


Рис. 2 – Интерфейс разрабатываемого ПО

Модуль визуализации отвечает за отрисовку данных в виде графиков/гистограмм в режиме реального времени. В качестве этого модуля был разработан и опубликован под свободной лицензией авторский построитель диаграмм, отвечающий как необходимым требованиям по времени, так и современным потребностям визуального оформления.

ПО написано на языке программирования C++ с использованием библиотеки Qt и в на-

стоящее время скомпилировано для двух платформ – GNU/Linux и MS Windows.

II. ОБЗОР СУЩЕСТВУЮЩИХ СРЕДСТВ ПОСТРОЕНИЯ ГРАФИКОВ В QT

При необходимости реализовать функционал построения графиков Qt-разработчик сталкивается с выбором из нескольких вариантов, причем это одна из немногочисленных задач, где на сегодняшний день все ещё сохраняется конкурентоспособность проприетарных Qt-компонентов. Так, к актуальным строителям графиков для Qt можно отнести следующие:

- QCustomPlot [2];
- QChart [3];
- Qwt [4];
- ChartDirector [5].

Однако, ни один из перечисленных строителей не отвечает поставленным требованиям.

III. ОСОБЕННОСТИ ПРОГРАММНОЙ РЕАЛИЗАЦИИ ОТРИСОВЩИКА

Причиной разработки собственного инструмента построения графиков для Qt является недостаточная производительность одних бесплатных фреймворков, которая делает их мало пригодными для отрисовки динамически изменяющихся графиков в режиме реального времени, и сильно ограниченные визуальные возможности других. Для написания виджета отрисовки графиков выбор между языком C++ и языком разметки QML был сделан в пользу C++ в силу требований максимальной производительности кода.

Программная реализация компонента Plotter представляет собой многоуровневую архитектуру, что позволяет настроить строителя под требования пользователя.

В общем случае структура классов принимает вид, изображенный на рисунке 3.

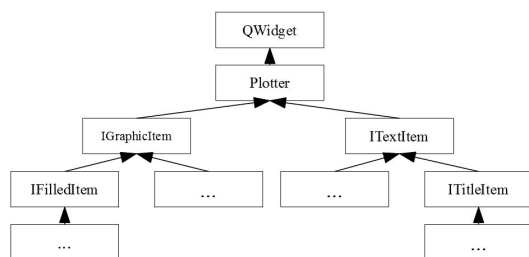


Рис. 3 – Структура классов отрисовщика

Как видно из данной диаграммы, основные интерфейсные базовые классы включают IGraphicItem, IFilledItem, ITextItem, ITitleItem.

Подклассы, унаследованные непосредственно от IGraphicItem соответствуют различным видам линий: для отрисовки главной и второстепенной сеток, для отрисовки нулевых линий, для отрисовки штрихов главной и второстепенной сеток.

Подклассы интерфейса IFilledItem отвечают за отрисовку графических элементов с заливкой: внешняя область виджета, заливка области построения, заливка области под графиком.

Подклассы интерфейса ITextItem отвечают за отрисовку текстовых полей: подписи осей, названия осей, легенда, название графика.

Также стоит упомянуть, что благодаря многоуровневой архитектуре расчеты для отрисовки графических элементов отделены от непосредственной отрисовки, а в качестве дополнительных бонусов реализован экспорт графиков в формат масштабируемой векторной графики SVG, и опциональное хранение пакета настроек графика в конфигурационном файле.

На рисунке 4 приведен минимальный код для подключения, инициализации, добавления данных и отрисовки графика.

```

1 Plotter* plotter = new Plotter;
2 const QString chart_name = "Chart 1";
3 plotter->add_chart(chart_name);
4 const QPen pen = QPen(QColor(220,170,170), 1, Qt::SolidLine);
5 plotter->chart(chart_name)->set_pen(pen);
6 plotter->chart(chart_name)->add_data(x, y);
7 plotter->scroll_graph();
8 plotter->replot();
  
```

Рис. 4 – Минимальный код для работы с отрисовщиком

Строка 3 отвечает за добавление графика на область построения. Строка 5 отвечает за инициализацию пера отрисовки. Строка 6 отвечает за добавление данных. Строка 7 необходима для прокрутки области построения графика. Строка 8 инициализирует общее обновление для области построения.

Пример работы строителя приведен на рисунке 5.

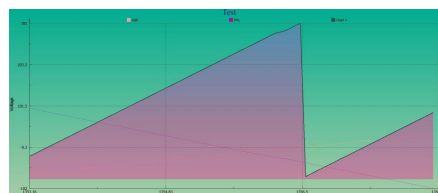


Рис. 5 – Пример работы строителя диаграмм

1. Латий, О. О. Микроконтроллерная система для биометрической оценки состояния пользователя ПК // Д. А. Костюк, О. О. Латий, В. П. Шамонин // Информационные технологии и системы 2017 (ИТС 2017) : материалы международной научной конференции, Минск, БГУИР, 25 октября 2017 г. – Минск, 2017. – С. 238–239.
2. QCustomPlot Qt Plotting Widget QCustomPlot - Introduction [Electronic resource] – Mode of access: www.qcustomplot.com/. – Date of access: 14.09.2018.
3. QChart Class [Electronic resource] – Mode of access: <https://doc.qt.io/qt-5.10/qchart.htm>. – Date of access: 14.09.2018.
4. Qwt User's Guide: Qwt - Qt Widgets for Technical Applications [Electronic resource] – Mode of access: qwt.sourceforge.net. – Date of access: 14.09.2018.
5. ChartDirector Chart Component and Control Library [Electronic resource] – Mode of access: www.advsofteng.com – Date of access: 14.09.2018.

Секция «Математическое
моделирование и компьютерная
графика»

Section «Mathematical modeling and
computer graphics»

СРАВНИТЕЛЬНОЕ МОДЕЛИРОВАНИЕ МАРКОВСКИХ ЦЕПЕЙ В СРЕДЕ ПРОГРАММИРОВАНИЯ R И MICROSOFT EXCEL

Аникин В. И., Козырин Д. И.

Кафедра «Информационный и электронный сервис», Поволжский государственный университет сервиса,
Кафедра «Информационный и электронный сервис», Поволжский государственный университет сервиса
Тольятти, Россия

E-mail: anikin_vi@mail.ru, kozyrin-di@yandex.ru

В Microsoft Excel и R построена табличная модель для преобразования матрицы переходов поглощающей марковской цепи к канонической форме с произвольной размерностью для сравнения сред программирования и выявления их недостатков и преимуществ.

ВВЕДЕНИЕ

Перед началом создания модели всегда встает вопрос: “В какой программной среде создавать модель?” В данной статье будет рассмотрены две программные среды для моделирования марковских цепей: R и Microsoft Excel. Цепь Маркова являются мощным и широко используемым инструментом для анализа множества стохастических (вероятностных) систем во времени [1]. Программа Excel все еще является лидером на рынке программ обработки электронных таблиц, именно она определяет тенденции развития в этой области. Excel позволяет быстро анализировать большие массивы данных. В Excel имеется более 400 математических, статистических, финансовых и других функций, позволяющих строить большие иерархические структуры и обширные модели. Но в настоящее время язык R также является одним из ведущих статистических инструментов в мире. Он активно применяется в молекулярной биологии, науках об окружающей среде, генетике, биоинформатике.

R является доступным и хорошо сбалансированным представлением теории случайных процессов с акцентом на реальных применениях теории вероятностей в естественных и социальных науках. Использование имитации с помощью популярного статистического программного обеспечения R подтверждает теоретические выводы при помощи практических демонстраций [2]. Язык R имеет широкий спектр статистических и численных методов и обладает хорошей расширяемостью с помощью пакетов, на данный момент их насчитывается порядка 12500.

В статье будет рассмотрен один пример разработки в среде Excel и 2 примера в R, так как подробная разработка в Excel были рассмотрены нами в прошлых материалах [3] и [4].

Цель статьи: Разработка моделей марковских цепей в среде R и Excel, сравнение по удобству написания кода и визуализации модели, нахождение общих достоинств и недостатков этих сред программирования.

Для оценки сред программирования была использована готовая модель на Excel из нашей опубликованной статьи [4] (рис. 1).

Рис. 1 – Модель марковской цепи в Microsoft Excel

А именно будут рассмотрены формулы таблиц I-Q, N, NR.

I. СРЕДА РАЗРАБОТКИ MICROSOFT EXCEL

В VBA Excel сначала нужно выделить редактируемый диапазон ячеек и только затем вносить в них формулы, поэтому формула становится длиннее.

I-Q: формула в табличной среде: =R[-4]C-R[-8]C
Формула в VBA:ActiveSheet.Range(Cells(x-p+x-p+13, 5+x), Cells(x+x+x-p-p+12, x+x-p+4)).Select

Selection.FormulaR1C1 = -R[-"& x-p+1 & "]C-R[-"& x-p+x-p+2 & "]C"

N: формула в табличной среде: МОБР(R[-4]C:R[-2]C[2])

Формула в VBA:ActiveSheet.Range(Cells(x-p+x-p+x-p+14, 5+x), Cells(x+x+x+x-p-p-p+13, x+x-p+4)).Select

Selection.FormulaArray = -MINVERSE(R[-"& x-p+1 & "]C:R[-2]C["& x-p-1 & "]")

NR: формула в табличной среде: =МУМНОЖ(RC[-10]:R[2]C[-8];RC[-3]:R[2]C[-2])

Формула в VBA:ActiveSheet.Range(Cells(x+x+x-x-p-p-p+16, x-p+8+p), Cells(x+x+x+x-x-p-p-p-p+15, x-p+p+p+7)).Select

Selection.FormulaArray = -MMULT(RC[-"& x-p+p+5 & "]R["& x-p-1 & "]C[-"& p+6 & "],RC[-"& p+1 & "]R["& x-p-1 & "]C[-"& 2 & "]")

Минусом этой среды является добавление и удаление диапазонов настолько увеличивает время написание кода, что при разработке табличной модели приходится выбирать требуемые размерности заранее и планировать наперед размерность каждой отдельной таблицы, это тема была рассмотрена в прошлой статье по построению итерационной модели марковских цепей в Microsoft Excel с произвольной размерностью матрицы переходов [3].

Преимущества среды разработки: всё, с чем работает пользователь, содержится в едином пространстве таблицы: пометки и комментарии, графики и диаграммы, входные, промежуточные и выходные данные, документация. Поэтому написание алгоритмов в табличном процессоре не требует специальных знаний у пользователя в области программирования.

Недостаток среды разработки: пользователь пишет код программы, используя низкоуровневые объекты — ячейки, привязанные к сетке. Табличным процессор использует только ячейки и формулы, средств абстракции более высокого уровня нет.

II. СРЕДА РАЗРАБОТКИ R

Программирование в R имеет более привычный вид и математические формулы пишутся короче, но для наглядности был использован графический интерфейс для вывода результатов (рис. 2).

Пример 1. Листинг:

```
k <- matrix(c(0,0.5,0,0.5,0,0.5,0,0,0.5,0,0,0.5,0,0,0.5,0,0,0,0,1), nrow=5, ncol=5, byrow=TRUE)
q <- matrix(c(0,0.5,0,0.5,0,0.5,0,0.5,0), nrow = 3, ncol = 3, byrow = TRUE)
i <- matrix(c(1,0,0,0,1,0,0,0,1), nrow = 3, ncol = 3, byrow = TRUE)
r <- matrix(c(0.5,0,0,0,0.5), nrow = 3, ncol = 2, byrow = TRUE)
h <- i-q # I-Q
n <- solve(h) # N
nr <- n
```

I-Q:					N:					NR:				
	i	o	p			1	2	3			1	2	3	
i	1.0	-0.5	0.0		1	1.5	1	0.5		1	0.75	0.25		
o	-0.5	1.0	-0.5		2	1.0	2	1.0		2	0.50	0.50		
p	0.0	-0.5	1.0		3	0.5	1	1.5		3	0.25	0.75		

Рис. 2 – Графический вывод таблиц в R

Написание кода модели в среде R происходит быстро и интуитивно понятно все результаты выводятся в консоль, но для визуальной проверки выходных данных нужно использовать графический интерфейс, который использовался для вывода таблиц (рис. 2).

Пример 2. Найти стационарное распределение погоды Марковской цепи (рис. 3).

$$P = \begin{matrix} & \begin{matrix} \text{Дождь} & \text{Снег} & \text{Ясно} \end{matrix} \\ \begin{matrix} \text{Дождь} \\ \text{Снег} \\ \text{Ясно} \end{matrix} & \begin{pmatrix} 1/5 & 3/5 & 1/5 \\ 1/10 & 4/5 & 1/10 \\ 1/10 & 3/5 & 3/10 \end{pmatrix} \end{matrix}$$

Рис. 3 – Матрица переходов

Листинг:

```
stationary <- function(mat)
x = eigen(t(mat))$vectors[,1]
as.double(x/sum(x))
mat <- matrix(c(1/5,3/5,1/5,1/10,
4/5,1/10,1/10,3/5,3/10),nrow=3, byrow=T)
> mat
0.2 0.6 0.2
0.1 0.8 0.1
0.1 0.6 0.3
lambda <- stationary(mat)
lambda %% mat
0.1111111 0.75 0.1388889
```

Преимущества среды разработки: большая библиотека статистических методов, имеется большая база дополнительных пакетов для скачивания, в языке R можно взаимодействовать с другими языками программирования, работать можно с различными форматами данных.

Недостаток среды разработки: сложность в использовании для начинающих пользователей, мало документации на русском языке.

III. ВЫВОДЫ

Microsoft Excel и R имеют очень обширный инструментарий для разработки моделей, каждая имеет свои недостатки и преимущества. Если у пользователя небольшой опыт в программировании, то ему хорошо подойдет Microsoft Excel, он дает возможность строить большие модели и визуализировать результаты моделирования, что позволяет даже сторонним людям объяснить принцип работы модели шаг за шагом. Если же нужно работать с данными, часто изменяющими размерность, или использовать для работы средства абстракции более высокого уровня, то лучше подойдет R. Но самым разумным представляется использование этих великолепных программных сред в тандеме, обмениваясь данными между ними посредством csv-файлов или разделяемой базы данных.

IV. СПИСОК ЛИТЕРАТУРЫ

- Ching, W.-K. Markov Chains: Models, Algorithms and Applications / W.-K. Ching, M. K. Ng // New York: Springer Science + Business Media. – 2006. – 205 p.
- Dobrow Robert P. Introduction to Stochastic Processes with R / Robert P. Dobrow // Hoboken: Wiley, – 2016. – 504 p.
- Аникин В. И., Аникина О. В., Козырин Д. И. Итерационные модели марковских цепей в Microsoft Excel / В. И. Аникин. О. В. Аникина. Д. И. Козырин. // Информационные технологии и системы 2017 (ИТС 2017). – Минск: БГУИР, – 2017. – с. 172-173.
- Аникин В. И., Козырин Д. И. Преобразование матрицы переходов поглощающей марковской цепи к канонической форме в Microsoft Excel / В. И. Аникин. О. В. Аникина. Д. И. Козырин. // Синергетика природных, технических и социально-экономических систем: – Изд-во ПВГУС, – 2018. – 74-79 с.

ОБРАЗОВАТЕЛЬНАЯ ИГРА «КОТЁНОК ЗВУКОВКИН»

Ключникова К. С., Бычков П. А.

ГУО «Ясли-сад №2 агр.Буйничи» Могилёвской области Могилёвского района
Кафедра вычислительных методов и программирования, Белорусский государственный университет
информатики и радиоэлектроники
Могилёв, Республика Беларусь; Минск, Республика Беларусь
E-mail: k-klyuchnikova.93@mail.ru, pashaby4ok@mail.ru

Рассматриваются актуальность, цель, содержание и задачи образовательной игры «Котёнок Звуковкин», а так же рассматриваются существующие аналоги разрабатываемого проекта.

ВВЕДЕНИЕ

Данный проект объединяет в себе разнообразные творческие задания и упражнения по закреплению правильного произношения, что актуально и необходимо, так как методика автоматизации звуков различных категорий однотипна и лишь требует использования различного речевого материала, насыщенного тем или иным звуком. Кроме этого, в настоящее время речевые расстройства усложнились повсеместным нарушением саморегуляции у детей, и, как следствие, их неспособностью удерживать, автоматизировать и самостоятельно вводить в речь уже поставленные звуки. Уже первый этап постановки звука затрудняется из-за низкой способности ребенка концентрироваться, понимать и выполнять словесную инструкцию. Ребенку бывает трудно даже посмотреть на себя в зеркало и удерживать внимание на несколько секунд. Использование данного проекта для закрепления поставленных звуков и введение их в речь позволяет сделать этот процесс более быстрым, а также повысить интерес детей к логопедическим занятиям.

I. ОБЗОР СУЩЕСТВУЮЩИХ АНАЛОГОВ

В данном разделе описываются существующие аналоги разрабатываемого проекта.

- Приложение «Логопед» в App Store предлагает пользователю самостоятельно провести логопедическое экспресс-тестирование и выявить имеющиеся у ребенка проблемы в произношении звуков. По результатам тестирования пользователь получит рекомендации и комплекс упражнений, составленный с учетом речевых проблем ребенка и ориентированный на их преодоление. По заверениям разработчика тесты и упражнения составлены профессиональными врачами-логопедами. Стоимость приложения - 149 российских рублей.
- Приложение «Артикуляция Логопедия» в GooglePlay предлагает более 1200 слов и изображения для практики логопедии и распознавание голоса в режиме реального времени.

Приложение бесплатное, рейтинг - 3,7 из 5.

- Приложение «Домашний логопед» в GooglePlay предлагает 350 уроков для отработки проблемных звуков, красивые рисунки к каждому упражнению, игры, загадки, пословицы, поговорки и речевую, дыхательную гимнастику, а так же дополнительный материал для родителей по проблеме неречевого и фонетического слуха. Приложение бесплатное, рейтинг - 4,5 из 5.

II. ЦЕЛЬ И ЗАДАЧИ ПРОЕКТА

Целью данного проекта является разработка игры для автоматизации звуков с учётом индивидуальных особенностей каждого ребенка, что позволит ускорить процесс преодоления речевых нарушений и повысить мотивацию детей к преодолению дефектов речи.

Задачи:

- Подобрать приемы и методы по закреплению речевых навыков, позволяющих облегчить процесс автоматизации звуков у дошкольников с неярко выраженным общим недоразвитием речи.
- Создавать игровые пособия для логопедических занятий, как способ формирования правильной речи и коррекции её недостатков у детей дошкольного возраста.
- Повысить эффективность образовательного процесса.

III. НОВИЗНА И ИННОВАЦИОННОСТЬ

Использование в коррекционной работе компьютерных игр, приложений поддерживает у детей с различной речевой патологией познавательную активность, повышает эффективность логопедической работы в целом. Их применение интересно, познавательно и увлекательно детям.

Подобные проекты позволяют работать с любыми речевыми единицами от звука до текста, решать разнообразные задачи, внести игровые моменты в процесс коррекции речевых нарушений, многократно дублировать необходимый тип упражнений и речевой материал, использовать различный стимульный материал (картинки, буквы, слоги, слова, предложения, звучащую речь), работать на разных уровнях сложности

в зависимости от возможностей воспитанника, одновременно осуществляя коррекцию восприятия, внимания, памяти, мышления.

IV. СОДЕРЖАНИЕ

Данный проект представляет собой игру с сохранением этапов логопедического процесса, применяется на этапе закрепления звука в слогах и словах. В ходе игры решаются следующие задачи:

- развитие артикуляционной моторики;
- закрепление артикуляции и изолированно произношения звуков;
- совершенствование навыка слогового анализа слов;
- развитие логического мышления.

Игра начинается с сюрпризного момента — появления главного героя — котенка Звуковкина, что, конечно же, заинтересовывает ребенка, активизирует его внимание. Звуковкин живет в деревне в домике с четырьмя окошками. Первое окошко — артикуляционная гимнастика. Нажимая на второе/третье/четвертое окошко игрок выбирает необходимый обрабатываемый звук из группы звуков (свистящие, шипящие, сонорные). После чего будет переходить к списку уровней, которые будут открываться последовательно, после прохождения следующего. Также после прохождения уровня игроку будет предложено сразу перейти к следующему уровню либо выйти к списку уровней.

Рассмотрим данную игру на примере обрабатываемого шипящего звука [Ч]. Игра включает следующие задания: упр. «Паровозик» (см. рис. 1), «Волшебная дорожка», «Назови картинку», «Подбери слова на слоги», «Веселый поезд», «Четвертый лишний», «Поиск», «Закончи предложения».

Образовательная игра «Котёнок Звуковкин» создаётся на игровом движке Unity. Скрип-

ты для проекта написаны с использованием языка C#. Проект разрабатывается для смартфонов и планшетов, использующих операционную систему Android.

V. ВЫВОД

Какие же качества данного проекта являются дидактическими ценностями?

- способствует усвоению знаний не по необходимости, а по желанию детей и проходит не формально, а заинтересованно;
- помогает не только проявлять способности, но и совершенствовать их;
- является орудием обучения и воспитания, требуя от участников большого умственного напряжения и принося в тоже время большое удовольствие.

VI. СПИСОК ЛИТЕРАТУРЫ

1. 54-я научная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский Государственный Университет Информатики и Радиоэлектроники». Материалы конференции по направлению 2: Информационные технологии и управление. Секция «Вычислительные методы и программирование» / К. С. Ключникова, П. А. Бычков: Минск БГУИР, 2018 - 64 с. ;
2. Комарова, Л.А. Автоматизация звуков Ч, Щ в игровых упражнениях. Альбом дошкольника / Л.А. Комарова. — М.: Издательство ГНОМ и Д, 2009. — 32 с.;
3. Нищева, Н.В. Веселая артикуляционная гимнастика. Наглядно-дидактическое пособие / Н.В. Нищева. — СПб.: ООО «ИЗДАТЕЛЬСТВО «ДЕТСТВО-ПРЕСС», 2009;
4. Волкова, Л.С. Логопедия: учебник для студентов дефектол. фак. пед. вузов / Л.С. Волкова, С.Н. Шаховская. - М.: Гуманит. изд. центр ВЛАДОС, 1998. — 680 с.;
5. Филичева, Т.Б., Соболева А.Р. Развитие речи дошкольника: Методическое пособие с иллюстрациями / Т.Б. Филичева, А.Р. Соболева. - Екатеринбург: Изд-во «АРГО», 1996. - 80 с.



Рис. 1 – Упражнение «Паровозик»

АЛГОРИТМ АППРОКСИМАЦИИ ЧАСТОТНЫХ ХАРАКТЕРИСТИК СВЧ ТРАНЗИСТОРОВ, ПРЕДСТАВЛЕННЫХ В ЧИСЛЕННОМ ВИДЕ, НА ДИСКРЕТНОМ РЯДЕ ЧАСТОТ

Исаев В. О., Свириденко А. А., Шанчук Р. А.

Кафедра автоматики, радиолокации и приемо-передающих устройств, кафедра тактики и вооружения зенитных ракетных войск факультета противовоздушной обороны, Учреждение образования «Военная академия Республики Беларусь»
Минск, Республика Беларусь
E-mail: {ystasmoz, svirid2785}@gmail.com

Представлен алгоритм аппроксимации частотных характеристик современных СВЧ транзисторов, представленных в численном виде, что противоречит постановке задачи на широкополосное согласование аналитическими методами, которые требуют строгого описания объектов исследования. Это означает, что согласуемая нагрузка должна представляться в виде математической формулы. Настоящая работа направлена на решение данной проблемы.

ВВЕДЕНИЕ

В настоящее время во всем мире наблюдается стремительное развитие радиоэлектронных систем в диапазоне сверхвысоких частот (СВЧ): систем сотовой и радиорелейной связи, радионавигации и радиолокации, телевидения и т.д. Убедиться в этом нетрудно, обратившись к широкому спектру самой передовой продукции в диапазоне СВЧ, выпускаемой ведущими корпорациями в этом секторе: TriQuint Semiconductor, Hittite Microwave Corporation, Excelics, RFMD, Mimix Broadband и рядом других. Компоненты, в частности СВЧ транзисторы, выпускаемые этими корпорациями, как правило имеют технический паспорт – «Datasheet», в котором указывается основная информация о рабочих параметрах, режимах работы и характеристиках транзистора. Частью этой информации являются заданные на дискретном ряде частот значения (модуль и фаза) элементов матрицы рассеяния. При проектировании СВЧ радиоэлектронных устройств (РЭУ) (СВЧ транзисторные усилители, преобразователи и умножители частоты, активные фильтры, антенные устройства и др.) важное значение имеет решение задач широкополосного согласования. При задании параметров рассеивания согласуемых СВЧ устройств в виде численных дискретных зависимостей модуля и аргумента от частоты задача согласования может быть решена исключительно численными методами. По-иному обстоит дело, когда согласующая цепь находится аналитическими методами. Здесь успех в решении задач согласования напрямую связан с определением адекватных математических моделей (дробно-рациональных функций) согласуемых нагрузок. Таким образом, актуальным является вопрос о нахождении математических моделей СВЧ РЭУ, представленных в численном виде на дискретном ряде частот, которые бы описывали параметры этих устройств с требуемой точностью. Это позволит

применять современные аналитические методики широкополосного согласования и послужит толчком для их дальнейшего развития.

МАТРИЦА РАССЕИВАНИЯ ЧЕТЫРЕХПОЛЮСНИКОВ

Для пассивного линейного четырехполосника (ЧП), включенного в СВЧ тракт с волновым сопротивлением, можно записать уравнения, определяющие линейную связь между падающими и отраженными волнами на входе и выходе ЧП, в виде:

$$U_{отр1} = S_{11}U_{пад1} + S_{12}U_{пад2}, \quad (1)$$

$$U_{отр2} = S_{21}U_{пад1} + S_{22}U_{пад2}. \quad (2)$$

Матрицу [S] называют матрицей рассеяния. Для ЧП эта матрица имеет размер . Она устанавливает связь между комплексными нормированными амплитудами отраженных и падающих волн в плечах ЧП. В матричной записи уравнения (1,2) приобретают вид:

$$\begin{bmatrix} U_{отр1} \\ U_{отр2} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \times \begin{bmatrix} U_{пад1} \\ U_{пад2} \end{bmatrix}. \quad (3)$$

Элементы волновой матрицы рассеяния имеют ясный физический смысл и могут быть измерены сравнительно простым способом, в частности с помощью измерительной линии. При работе СВЧ – четырехполосника на согласованную нагрузку отраженная волна, на его выходе, отсутствует, а из соотношения (3) следует:

$$S_{11} = \frac{U_{отр1}}{U_{пад1}},$$

$$S_{21} = \frac{U_{пад2}}{U_{пад1}}.$$

где S_{11} – комплексный коэффициент отражения от входа исследуемого ЧП, а S_{21} –

комплексный коэффициент передачи ЧП. В общем случае он учитывает как активные потери в четырехполюснике, так и потери на отражение. Элементы S_{22} и S_{12} имеют аналогичный смысл, но соответствуют обратному включению ЧП (при этом выход ЧП соединяют с генератором, а на вход его включают согласованную нагрузку). Значения матрицы рассеяния описывают свойства ЧП лишь на заданной частоте. Для представления ЧП в полосе частот элементы матриц рассеяния преобразуются в рациональную функцию вида:

$$f(i\omega) = \frac{a_0 + a_1(i\omega) + \dots + a_k(i\omega)^k}{b_0 + b_1(i\omega) + \dots + b_n(i\omega)^n}. \quad (4)$$

I. АЛГОРИТМ АППРОКСИМАЦИИ ЧАСТОТНЫХ ХАРАКТЕРИСТИК СВЧ ТРАНЗИСТОРОВ, ПРЕДСТАВЛЕННЫХ В ЧИСЛЕННОМ ВИДЕ, НА ДИСКРЕТНОМ РЯДЕ ЧАСТОТ

Взяв за основу дробно-рациональную функцию (4) с неизвестными коэффициентами при переменной s можно максимально точно аппроксимировать заданные в табличном виде модуль и фазу коэффициента отражения S_{11} СВЧ устройства. Так как рассматриваемые модуль и фаза коэффициента отражения S_{11} являются комплексными, то для поиска функции, максимально точно описывающей транзистор с заданными параметрами, необходимо воспользоваться некоторыми свойствами комплексных чисел. Как известно, модуль комплексного числа S_{11} можно представить в виде:

$$S_{11} = |S_{11}| = \sqrt{A^2 + B^2}; \quad (5)$$

$$\varphi = \arctg \frac{B}{A}; \quad (6)$$

где A – действительная часть комплексного элемента матрицы рассеяния СВЧ устройства, а B – его мнимая часть. Представив функцию (4) через четные и нечетные части ее числителя и знаменателя [1]:

$$f(i\omega) = \frac{P(i\omega)}{Q(i\omega)} = \frac{(m_1 + n_1)}{(m_2 + n_2)}.$$

Умножим $P(i\omega)$ и $Q(i\omega)$ на $(m_2 - n_2) = Q(-i\omega)$:

$$\begin{aligned} f(i\omega) &= \frac{P(i\omega)}{Q(i\omega)} = \frac{(m_1 + n_1) m_2 - n_2}{(m_2 + n_2) m_2 - n_2} \\ &= \frac{(m_1 m_2 - n_1 n_2) + (n_1 m_2 - m_1 n_2)}{m_2^2 - n_2^2}. \end{aligned}$$

где $m_1 = a_0 + a_2(i\omega)^2$ – четная часть числителя функции $f(i\omega)$; $m_2 = b_0 + b_2(i\omega)^2$ – четная часть знаменателя функции $f(i\omega)$; $n_1 = a_1(i\omega)$ – нечетная часть числителя функции $f(i\omega)$; $n_2 = b_1(i\omega)$ – нечетная часть знаменателя функции $f(i\omega)$. Из теории цепей известно [1], что действительная

и мнимая части дробно-рациональной функции (4) находятся как:

$$\mathbf{Re}f(i\omega) = \frac{(m_1 m_2 - n_1 n_2)}{m_2^2 - n_2^2};$$

$$\mathbf{Im}f(i\omega) = \frac{(n_1 m_2 - m_1 n_2)}{m_2^2 - n_2^2}.$$

В этом случае, действительную и мнимую части требуемой функции можно представить как:

$$\mathbf{Re}f(i\omega) = \frac{(a_0 + a_2(i\omega)^2)(b_0 + b_2(i\omega)^2) - a_1 b_1(i\omega)^2}{(b_0 + b_2(i\omega)^2)^2 - (b_1(i\omega))^2},$$

$$\mathbf{Im}f(i\omega) = \frac{(b_1(i\omega))(b_0 + b_2(i\omega)^2) - (a_0 + a_2(i\omega)^2)b_1(i\omega)}{(b_0 + b_2(i\omega)^2)^2 - (b_1(i\omega))^2},$$

Исходя из (5) и (6) модуль и фаза примет вид:

$$|\rho(i\omega)| = \sqrt{(\mathbf{Re}f(i\omega))^2 + (\mathbf{Im}f(i\omega))^2}$$

$$\eta(i\omega) = \arctg \left(\frac{(b_1(i\omega))(b_0 + b_2(i\omega)^2) - (a_0 + a_2(i\omega)^2)b_1(i\omega)}{(a_0 + a_2(i\omega)^2)(b_0 + b_2(i\omega)^2) - a_1 b_1(i\omega)^2} \right)$$

Используя численный метод решения задачи Чебышевской аппроксимации и выбрав интервал частот, требуемый для согласования транзистора, получаем системы неравенств [2]:

$$\begin{cases} |S_{11}(\omega_{min}) - \rho(\omega_{min})| \leq \delta_{\rho(\omega_{min})}; \\ |S_{11}(\omega_1) - \rho(\omega_1)| \leq \delta_{\rho(\omega_1)}; \\ \dots; \\ |S_{11}(\omega_{max}) - \rho(\omega_{max})| \leq \delta_{\rho(\omega_{max})}. \end{cases}$$

$$\begin{cases} |\varphi(\omega_{min}) - \eta(\omega_{min})| \leq \delta_{\eta(\omega_{min})}; \\ |\varphi(\omega_1) - \eta(\omega_1)| \leq \delta_{\eta(\omega_1)}; \\ \dots; \\ |\varphi(\omega_{max}) - \eta(\omega_{max})| \leq \delta_{\eta(\omega_{max})}. \end{cases}$$

В качестве целевой функции выбран параметр δ , который будем минимизировать путём подбора коэффициентов $a_0, a_1, a_2, \dots, a_k$ и $b_0, b_1, b_2, \dots, b_n$. Для упрощения поиска требуемой функции так же примем во внимание следующее: ошибка всей функции будет тем больше, чем больше сумма каждой из ошибок по отдельности, т.е.

$$\begin{cases} |\delta_{\rho(\omega_{min})} + \delta_{\eta(\omega_{min})}| = \delta; \\ |\delta_{\rho(\omega_1)} + \delta_{\eta(\omega_1)}| = \delta; \\ \dots; \\ |\delta_{\rho(\omega_{max})} + \delta_{\eta(\omega_{max})}| = \delta. \end{cases}$$

Следовательно, искомая функции будет та, у которой δ будет минимальной.

1. Карни, Ш., Теория цепей. Анализ и синтез. – М. «Связь», 1973. – 269с.
2. Белецкий, А.Ф. Теория линейных электрических цепей / А.Ф. Белецкий. – М.: Радио и связь, 1986. – 544.

РАСЧЕТ ТЕРМОУПРУГИХ ДЕФОРМАЦИЙ В СЭНДВИЧ-ПАНЕЛИ

Курочка К. С., Комракова Е. В.

Кафедра Информационных технологий, Гомельский государственный технический университет имени П.О.Сухого

Гомель, Республика Беларусь

E-mail: kurochka@gstu.by, 5366253@mail.ru

На основе метода конечных элементов проведено численное исследование распределения температур и термоупругих деформаций в сэндвич-панели. Численный алгоритм разработан на основе теории термоупругости Грина-Линдсея с учетом времени тепловой релаксации. Рассмотрено термоупругое деформирование поперечно нагруженной упругой сэндвич-панели

ВВЕДЕНИЕ

Для твердых тел изменение внешней температуры вызывает внутренние механические напряжения. Этот эффект особенно выражен в случае неравномерного нагрева. В ряде случаев возникающие при этом деформации могут быть сопоставимы или больше деформаций, вызываемых действием внешней нагрузки.

Конструктивные элементы типа сэндвич-панелей, состоящие из прочных и жестких внешних слоев, низкопрочного и легкого внутреннего слоя широко используются в различных отраслях техники, так как для них характерны высокая прочность и жесткость, небольшая средняя плотность, хорошая способностью поглощать энергию при ударных воздействиях. Кроме того, трехслойные конструктивные элементы обеспечивают хорошие звуко- и теплоизолирующие свойства, а также обладают высокой технологичностью и вибростойкостью.

В ряде случаев составные части этих конструктивных элементов подвергаются стационарному равномерному тепловому воздействию, при котором изменяются физико-механические свойства материалов слоев и возникают градиенты температуры, сопровождающиеся неодинаковым тепловым изменением размеров частей элементов. Знание величины и характера действия тепловых напряжений необходимо для всестороннего анализа прочности конструкции. При этом необходимо учитывать, что тепловые напряжения в конструктивных элементах, особенно в сочетании с механическими напряжениями от внешних сил, могут привести к полному разрушению композитных тонкостенных конструкций и термовспучиванию. Кроме того, сэндвич-панели обладают следующими специфическими особенностями – резко выраженной анизотропией их механических характеристик, ослабленным сопротивлением поперечным деформациям, существенным различием механических и теплофизических характеристик слоев. Эти факторы имеют принципиальное значение, как отмечается в [1], при расчете деформаций и температур.

I. ПОСТАНОВКА ЗАДАЧИ

Объектом численного исследования являлась сэндвич-панель (рисунок 1) нагружаемая распределенными по верхнему слою механическими и тепловыми воздействиями (при этом предполагалось, что в нулевой момент времени как перемещения, так и приращение температуры были равны нулю). Для упрощения расчета, ввиду симметричности рассматриваемой панели, расчет производился для 1/4 части панели. На отсекаемых частях расчетной модели были заданы условия зеркальной симметрии.

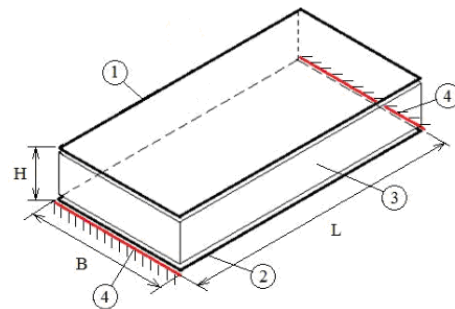


Рис. 1 – Схема нагружения сэндвич-панели: 1 - верхний металлический слой; 2 - нижний металлический слой; 3 - внутренний слой; 4 - места закрепления

Термоупругий изгиб пластины, описывается системой уравнений[2]:

$$\rho \frac{\partial^2 u}{\partial t^2} = \mu \Delta u + (\lambda + \mu) \text{grad div} u - \gamma \text{grad} T + \rho F, \quad (1)$$

$$\frac{\partial T}{\partial t} + \gamma \frac{T}{c\rho} \frac{\partial}{\partial t} \text{div} u = \alpha \Delta T, \quad (2)$$

где u , T – деформации и температуры в рассматриваемой точке пространства.

Для учета температурной зависимости коэффициента теплопроводности материалов сэндвич-панелей будем использовать теорию термоупругости Грина-Линдсея[3] для этого введем в расчет время тепловой релаксации. Данный параметр введем для того, чтобы устранить бесконечную скорость распространения тепловых волн. Данная теория не нарушается классический закон Фурье.

Уравнение состояния для теории обобщенной термоупругости с одним временем релаксации записывается в виде

$$\sigma_{ij} = \lambda e \delta_{ij} + 2\mu e_{ij} - \gamma(T - T_0)\delta_{ij}. \quad (3)$$

Уравнение теплопроводности, соотношение между напряжениями и перемещениями соответственно имеют вид

$$kT_{ii} = \rho C_E(\dot{T} + \tau_0 \ddot{T}) + \gamma T_0(e_{i,j} + \tau_0 \ddot{e}_{i,j}) - (Q - \tau_0 \ddot{Q}); \quad (4)$$

$$e_{ij} = (u_{i,j} + u_{j,i})/2; \quad (5)$$

где λ, μ – постоянные Ламе; ρ – плотность материала; C_E – удельная теплоемкость при постоянном напряжении; Q – мощность источника тепла на единицу массы; T – абсолютная температура; $T_0 = \delta_0 \rho c_0^2 / (\gamma_0 E_0) = (\delta_0 / \alpha_T)(1 - \nu) / (1 + \nu)$ – исходная температура; σ_{ij} – компоненты тензора напряжений; $e = \epsilon_{ii}$; ϵ_{ii} – компоненты тензора деформации; u_i – компоненты вектора перемещения; k – теплопроводность; τ_0 – время релаксации[4].

В конечноэлементной формулировке уравнения могут быть переписаны в виде[5]:

$$[K]\{\delta\} + [C] \frac{\partial}{\partial t} \{\delta\} + [M] \frac{\partial^2}{\partial t^2} \{\delta\} + \{F\} = 0, \quad (6)$$

$$[C] \frac{\partial \{\Phi\}}{\partial t} + [K]\{\Phi\} = \{F\}, \quad (7)$$

где δ – деформации, а Φ – температуры в рассматриваемых точках пространства-времени.

II. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

При расчетах предполагалась, что ширина панели $B=0.4$ м, а длина панели $L = 0.8$ м. Материал верхнего и нижнего слоя сэндвич-панели – оцинкованная сталь толщиной 0.02 м, материал внутреннего слоя – пеноплекс толщиной 0.1 м. К верхнему слою панели прикладывалась комплексная нагрузка в виде равномерно распределённого давления 200 кПа) и теплового воздействия. Тепловое воздействие моделировалось путем контакта верхнего слоя сэндвич-панели с поверхностью. Расчеты проводились как с учетом физической зависимости свойств материалов слоев сэндвич-панели от температуры и механического напряжения, так и без этого учета. На основе предложенного подхода решена задача о нагреве по экспоненциальному временному закону верхней и нижней обкладок композитной

трехслойной пластины. Этот закон имеет следующий вид $T(t) = T_0 + \Delta T(1 - e^{-t/\tau})$.

Расчет проводился на основе метода конечных элементов [4]. Прикладная программа для определения напряженно-деформированного состояния сэндвич-панелей при больших приложенных нагрузках реализована на языке высокого уровня C#.

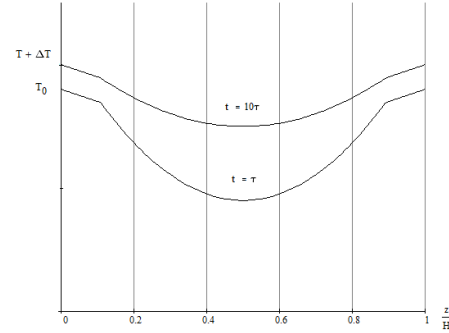


Рис. 2 – Зависимость температуры в центре пластины вдоль вертикали для двух значений t .

ЗАКЛЮЧЕНИЕ

В результате проведенных численных исследований получены распределение температуры и деформаций для разных точек сэндвич-панели в различные моменты времени. Из полученных результатов также следует, что при увеличении толщины внутреннего теплоизолирующего слоя, процесс распространения тепла, вдоль нормали к слоям сэндвич-панели, резко замедляется, что свидетельствует о правильности, с физической точки зрения, проведенных расчетов.

СПИСОК ЛИТЕРАТУРЫ

1. Chandrasekharaiah, D. S. Hyperbolic thermoelasticity a review of recent literature/ D. S. Chandrasekharaiah // Appl. Mech. Rev. 1998. V. 51, N 12. P. 705–729.
2. Самарский, А.А. Вычислительная теплопередача / А. А. Самарский, П. Н. Вабищев – М.: Едиториал УРСС, 2003 – 784 с.
3. Старовойтов, Э.И. Основы теории уругости, пластичности и вязкоупругости / Э. И. Старовойтов – Гомель: БелГУТ, 2001 – 344 с.
4. Liu, G.R. Finite Element Method A practical course, / G.R. Liu, S.S. Quek - PRENTICE-HALL, 2003 - 521 p.
5. Румянцев, А.В. Метод конечных элементов в задачах теплопроводности / А. В. Румянцев – Калининград, 2010 – 195 с.
6. Othman, M. I. A. Generalized thermo-microstretch elastic medium with temperature dependent properties for different theories/M. I. A. Othman, Kh. Lotfy , R. M. Farouk // Engng Anal. Boundary Element. 2010. V. 34. P. 229–237.

ЭРГОНОМИЧНЫЙ ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС КАК СРЕДСТВО ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ РАБОТЫ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ В УЧРЕЖДЕНИИ ВЫСШЕГО ОБРАЗОВАНИЯ

Нестеренков С. Н., Макаров М. И., Ющенко Н. В., Радкевич А. Д., Шатилова О. О., Рак Т. А.
Отдел информационных технологий центра информатизации и инновационных разработок БГУИР,
Кафедра ВМиП, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: nsn@bsuir.by, max.makarovv@gmail.com, nikitka.vit@gmail.com, arr.scout@gmail.com

Рассматривается автоматизированная система управления процессом получения высшего образования. Акцентируется внимание на важности разработки эргономичного пользовательского интерфейса в автоматизированной системе управления получения высшего образования. Выделяются ключевые аспекты построения эргономичных пользовательских интерфейсов.

ВВЕДЕНИЕ

Информационная подсистема управления современным университетом играет основополагающую роль и определяет эффективность и результативность как функций управления, так и всех остальных подсистем управления учебным заведением[1].

Интегрированная информационная система (ИИС) включает множество подсистем, таких как учет информации о студентах, группах, учебные планы специальностей, составление расписания и др.[2] Этими подсистемами пользуется множество сотрудников университета и они полагаются на ее качество, безотказность и удобство.

Разработка эргономичного пользовательского интерфейса для автоматизированной системы в учреждении высшего образования является одним из ключевых факторов для повышения качества работы системы, уменьшения количества допускаемых человеческих ошибок в бюрократическом аппарате и как следствие повышения удовлетворенности сотрудников своей работой и повышение эффективности учебного процесса.

I. ПРЕДПОСЫЛКИ РАЗРАБОТКИ ЭРГОНОМИЧНОГО ИНТЕРФЕЙСА ИИС

ИИС управления университетом разрабатывается, как и большинство других автоматизированных систем, на основе поставляемых задач диктуемых временем[3]. В частности, появляются новые требования по реализации той или иной функциональности, старые требования видоизменяются или исчезают, меняется структура подразделений предприятия и список обязанностей ее сотрудников. Все это в совокупности так или иначе меняет потребности и возможности ИИС.

С изменениями в системе нередко возрастает доля ошибок, падает скорость взаимодей-

ствия между человеком и машиной[4] в связи с тем, что человеку необходимо изучать новые или измененные возможности системы, а также при увеличении отклика самой системы на действия пользователя.

Система не является коммерческим продуктом, разрабатывается одним из подразделений университета, а значит тестирование на пользователях возможно только в рамках самого учреждения образования и отсутствует конкуренция, как на внешнем рынке, а значит у пользователя нет выбора и он обязуется использовать данную систему. Это поднимает вопросы удобства пользования ИИС.

II. ЭРГОНОМИЧЕСКИЕ ПАРАМЕТРЫ СИСТЕМЫ

Система представлена веб-сайтом в рамках учреждения образования, а значит существует ряд параметров по которым можно оценить юзабилити[4].

Одним из таких важных параметров является «восприимчивость», когда пользователь впервые зашел в систему, насколько ему будет все понятно, как легко он будет ориентироваться в структуре системы, сумеет ли понять какие действия ему необходимо производить для достижения своих целей.

Вторым ключевым показателем является «эффективность», как быстро у пользователя получится произвести действия в рамках ИИС для достижения результата.

Еще одним показателем является «запоминаемость». Насколько сократится время для достижения пользователем цели, которую он достигал ранее при повторном входе.

В ИИС университета существует множество пользователей с разными возможностями и обязанностями. Одним из ключевых факторов является то, что система предоставляет единый интерфейс для работы сотрудников, что упрощает взаимодействие между персоналом и подразделениями, повышает показатели запоминаемо-

сти и восприимчивости при смене обязанностей и должностей сотрудников.

Как и в любой другой системе человек-машина, в ИИС существует ряд проблем которые требуют решения. С точки зрения эффективности системы всегда встают новые вопросы при добавлении нового функционала в систему, изменении старого, повышении качества и эффективности существующего.

В связи со спецификой системы повышение эффективности взаимодействия в среде человек-машина достигается за счет сокращения как времени восприятия человеком элементов системы и количества действий для достижения поставленного результата, так и сокращения времени отклика системы на действия пользователя, минимизация или сведение к нулю количества ошибок генерируемых системой на корректные и некорректные обращения пользователя.

Для примера следует привести некоторые из ключевых точек где система может работать эффективнее. Одной из таких точек является генерация отчетов. Статистика системы показывает, что в некоторых случаях генерация длится несколько десятков секунд или даже более одной минуты. При генерации малого количества отчетов данного типа это может быть несущественно, но при частом обращении за данным функционалом пользователь испытывает неудобства. Другими точками увеличения эффективности могут быть подсистема работы с планами, где каждый план представлен в виде таблицы аналогичной бумажному плану, подсистема работы с личной карточкой студента и подсистема составления расписания где время отклика на некоторые действия составляет значительное время для пользователя, а также в ряде случаев возникают ошибки на действия человека.

В общем случае увеличение эффективности конкретного модуля влияет на пользователей которые имеют доступ к этой части системы.

III. СПОСОБЫ ПОВЫШЕНИЯ ЭРГОНОМИЧНОСТИ ИИС

Исходя из архитектуры и специфики ИИС, а также частей системы, где явно можно увеличить эффективность взаимодействия с пользователем, существует несколько подходов к решению данной проблемы.

Первый подход заключается в сокращении количества возможных ошибок, минимизации отклика между системой и пользователем, что в свою очередь затрагивает такие аспекты как правильно написанные запросы к базе данных, минимизация запросов между интерфейсом пользователя, представленным страницей в браузере, и сервером, минимизация времени обработки каждого запроса.

Вторым, не менее важным, подходом является качественное построение интерфейса пользователя[5],[6]. Он должен быть предельно понятен и удобен. Одним из ключевых факторов в данном случае является минимизация взаимодействия человека и системы. Элементы где присутствует возможность выбора, должны быть заполнены, если в качестве выбора предполагается всего 1 вариант. Элементы относящиеся к одному функционалу должны быть сгруппированы. Взаимодействие системы и человека должно быть информативным и однозначным, чтобы пользователь получал ожидаемый результат на свои действия. На более сложных шагах, или в случае неоднозначности выбора, должны быть предоставлены соответствующие подсказки.

С другой стороны пользовательский интерфейс должен быть как можно более простым при сохранении функциональности, не нагружать человека излишней информацией и не требовать от него предельного внимания и концентрации.

Как итог, ИИС должна быть модифицирована с учетом выполняемой сотрудниками университета работы, повышения удобства выполнения этой работы и эффективности взаимодействия человека с системой.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Нестеренков, С.Н. Основные принципы построения системы управления современным учреждением образования / С.Н. Нестеренков, О.О. Шатилова, Т.А. Рак // Дистанционное обучение - образовательная среда XXI века : материалы X Междунар. науч.-метод. конф., Минск, 7-8 декабря 2017 г. / Белорус. гос. ун-т информатики и радиозлектроники ; редкол.: Б.В. Никульшин [и др.]. - Минск, 2017. - С. 171.
2. Нестеренков, С. Н. Интегрированная информационная система как средство автоматизации управления образовательным процессом в учреждениях высшего образования / С. Н. Нестеренков, Т.А. Рак, О.О. Шатилова // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиозлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2017. - С. 212.
3. Шатилова, О.О. Современные подходы к созданию интегрированной информационной системы управления университетом / О.О. Шатилова, С.Н. Нестеренков, Т.А. Рак // Дистанционное обучение - образовательная среда XXI века : материалы X Междунар. науч.-метод. конф., Минск, 7-8 декабря 2017 г. / Белорус. гос. ун-т информатики и радиозлектроники ; редкол.: Б.В. Никульшин [и др.]. - Минск, 2017. - С. 172.
4. Вайнштейн Л.А., Эргономика: учебное пособие /Л.А. Вайнштейн.- Минск: ГИУСТ БГУ, 2010.-399 с.
5. Костин А.Н. Методологические подходы к юзабилити-проектированию // Человеческий фактор: проблемы психологии и эргономики. - 2013. - № 3. - с. 30-33.
6. Сергеев С.Ф. Юзабилити информационных систем в образовании: основные этапы юзабилити в тестировании // Образовательные технологии (г. Москва). - 2013. - № 2. - с. 57-63.

ПРОГРАММНОЕ СРЕДСТВО ДЛЯ МОДЕЛИРОВАНИЯ РАСПРОСТРАНЕНИЯ РАДИОНУКЛИДОВ В ОКРУЖАЮЩЕЙ СРЕДЕ НА ОСНОВЕ ГАУССОВОЙ МОДЕЛИ

Шамына А. Ю., Ардяко А. Д.

Факультет компьютерных систем и сетей, факультет информационных технологий и управления,
Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: kloop1996@gmail.com, ardyakon@gmail.com

Разработана система поддержки принятия экспертных решений в случае инцидента на радиационном объекте. В рамках данной работы в качестве такого объекта выступила БелАЭС. Разработанная система способна строить прогнозы для различных аварийных сценариев, причем каждый из них представляет собой набор данных, отражающих изотопный состав выброса и его активность. Результатом моделирования является полигонизированное покрытие, отражающее плотность радиоактивного загрязнения местности с географической привязкой. С учетом полученных значений плотности и определенных дозовых коэффициентов вычисляется величина дозовой нагрузки, в соответствии с которой система дает рекомендацию о проведении тех или иных профилактических или эвакуационных мероприятий.

ВВЕДЕНИЕ

В настоящее время в мире растет число потенциально опасных объектов, которые представляют угрозу возникновения аварий с радиационным загрязнением окружающей среды. Вместе с этим возникает необходимость создания программного комплекса, который бы мог осуществлять поддержку принятия экспертного решения в случае инцидента на радиационном объекте. В качестве объекта для моделирования была выбрана БелАЭС, первый энергоблок которой будет введен в эксплуатацию в 2019 году. Исходными данными для проведения численного моделирования выбраны аварийные сценарии, которые предусмотрены проектом БелАЭС, а также текущая метеобстановка. При создании программного средства учитывались особенности законодательства РБ в области обеспечения безопасности эксплуатации подобных объектов.

I. ОСНОВНЫЕ ПОЛОЖЕНИЯ МОДЕЛИ РАСЧЕТА

Основой для построения программной модели послужила гауссова модель распространения загрязняющих веществ в атмосфере, которая представляет собой набор эмпирических формул для описания многочисленных экспериментальных данных. Данная модель позволяет рассчитывать распространение загрязняющего вещества на расстояниях не более 10 км от источника и делать ориентировочную экспресс-оценку переноса на расстояниях не более 30 км. Т.е моделирование с использованием данной модели имеет локальный характер.

Концептуально в гауссовых моделях предполагается, что рассеивание в атмосфере неоседающего загрязняющего вещества по горизонтали и по вертикали происходит по нормальному закону распределения при постоянных направ-

лении и скорости ветра и условиях сохранения устойчивости атмосферы в течение времени переноса [1].

В обобщенном формализованном виде расчет приземной концентрации загрязняющего вещества в атмосфере согласно гауссовой модели осуществляется по формуле для источника конечного времени действия (1):

$$q(x, y, z, t) = M * F(x) * G_{con}(x, y, z) * t_s \quad (1)$$

где M – мощность источника выброса, Бк/с;

$F(x)$ – функция обеднения источника;

G_{con} – фактор разбавления для источника непрерывного времени действия;

t_s – время действия источника, с.

Следует отметить, что данная формула справедлива для нештилевых условий. При скорости ветра меньше 1 м/с расчет приземной концентрации осуществляется по другой методике. Значения приземной концентрации можно использовать для оценки ингаляционных доз облучения.

Используя значение приземной концентрации загрязняющего вещества в атмосфере можно произвести расчет плотности выпадения, обусловленного сухим и влажным фактором выведения. Расчет плотности выпадения в общем виде осуществляется по формуле (2):

$$D(x, y, t) = D_d(x, y, t) + D_w(x, y, t) \quad (2)$$

где $D_d(x, y, t)$ – плотность выпадения загрязняющего вещества на подстилающую поверхность, обусловленная сухим выведением;

$D_w(x, y, t)$ – плотность выпадения загрязняющего вещества на подстилающую поверхность, обусловленная влажным выведением.

Учет влажного выведения из атмосферы осуществляется только при наличии осадков в

момент аварии, причем каждый вид осадков имеет свою вымывающую способность.

Плотность выпадения рассчитывается точно вдоль оси следа (консервативно предполагается, что ось следа совпадает с направлением ветра) с дискретностью 250 м. Максимальное удаление точек расчета относительно оси следа выброса определяется значением горизонтальной дисперсии. На основе плотности выпадения загрязняющего вещества рассчитывается дозовая нагрузка на население.

II. КОНСТРУИРОВАНИЕ ПРОГРАММНОГО СРЕДСТВА

При разработке данного программного средства использовался язык программирования C#. Для построения графического интерфейса пользователя была использована технология WPF. Работа с картографическими слоями осуществляется с использованием библиотеки DotSpatial.

Первоочередной задачей при реализации данного программного средства является определение исходных данных для моделирования. Проектные аварийные сценарии определены в [2]. Для их хранения, а также хранения других статичных данных, необходимых для моделирования используется реляционная база данных.

При моделировании учитывается целый ряд фактических метеорологических параметров. На данный момент для их импорта используются открытые API погоды, однако возможно использование данных о погоде с АСКРО (автоматизированная система контроля радиационной обстановки) БелАЭС, а также близлежащих метеостанций для повышения точности расчета.

Для отображения полей концентраций выпадения загрязняющих веществ используется полигонизированное покрытие. Оно строится с использованием интерполяционного метода обратных радиусов, исходными точками для которого служат полученные в результате расчета точечные значения с заданной дискретностью. После чего данное покрытие привязывается к картографической подложке. Слои картографической подложки включают в себя контуры границ районов, областей и государственной границы РБ.

Также отображаются населенные пункты с населением от 100 человек и выше.

Учет влияния рельефа местности при моделировании осуществляется через коэффициент шероховатости подстилающей поверхности, который рассчитывается по методике, определенной в [3].

После получения значений плотности загрязнения для каждого радионуклида из isotopного перечня выброса, выполняется расчет мощности дозы внешнего облучения. Он формируется как сумма вкладов мощностей облучения от каждого дозообразующего радионуклида. На основании полученного значения мощности формируются рекомендации по проведению защитных мероприятий в соответствии с [4].

III. ЗАКЛЮЧЕНИЕ

В рамках создания данного программного средства была разработана программная реализация локальной модели распространения загрязняющих веществ в атмосфере, а также целого ряда вспомогательных модулей и компонентов, необходимых для вычисления параметров моделирования и анализа полученных результатов. Важной особенностью данной программной реализации модели атмосферной диффузии является возможность её использования не только в рамках моделирования последствий радиационных аварий, но и также других аварий с выходом неких загрязняющих веществ в атмосферу, т.к. фракционный состав выброса и его диффузные свойства параметризованы.

1. Методика расчета рассеяния загрязняющих веществ в атмосфере при аварийных выбросах. РД 52.18.717-2009. Обнинск, ООО «ПРИНТ-СЕРВИС» – 2009.
2. Белорусская АЭС. Блок 2. Предварительный отчет по обоснованию безопасности. Глава 15. Анализ аварий. Книга 7. БЛ-02065пм, ОАО «НИАЭП», 2013 г.
3. User's Guide for the Industrial Source Complex (ISC3) Dispersion Models. V. II – Description of Model Algorithms. EPA-454/B-95-003b. US Environmental Protection Agency, September 1995.
4. Санитарные нормы, правила и гигиенические нормативы «Гигиенические требования к проектированию и эксплуатации атомных станций», утвержденные постановлением Министерства здравоохранения Республики Беларусь № 39 от 30 марта 2010 г.

Секция «Системы обработки
информации»

Section «Systems of information
processing»

FUZZY SECURE SKETCH BIOMETRIC SCHEME BASED ON NON-BINARY TURBO CODES

Assanovich B. A., Veretilo Yu. N.

Information Systems and Technologies Department of Yanka Kupala State University of Grodno

National Anti-Doping Laboratory

Grodno, Minsk, Republic of Belarus

E-mail: bas@grsu.by, hullit.pakkard@gmail.com

The implementation of a fuzzy secure sketch biometric scheme for so-called fuzzy commitment using non-binary turbo codes with several times better performance and flexibility has been proposed.

INTRODUCTION

Recently, the implementation of reliable cryptographic systems based on fuzzy extractors (Fuzzy Extractor) using unreliable «noisy» biometric data is of particular interest in the literature. It is known that if in such systems the arising noise caused by the fuzziness of the biometric data is additive and leads to substitution errors, an effective solution is the use of noise-immune codes with large Hamming distance. One of the known approaches to creating such a system is to use a code-offset construction [1], which forms an auxiliary sketch (Secure Sketch) stored in the database. It is applied together with the error correcting code (n, k, d) and represents the offset D that «shifts» the code vector X of the applied noise-proof code containing the user's password S by the biometric measurement value B , i.e. $D = B - X$. In the subsequent biometric measurement B' , subtraction $D - B' = Y$, decoding Y and obtaining the password S' , generally coinciding with S , is performed.

To achieve the necessary efficiency (to minimize the probability of FAR and FRR), it is necessary to apply the «powerful» error-correction codes, for example BCH, increasing the Hamming distance to correct multiple errors [2], and also to non-binary noise-resistant codes (Reed-Solomon, Turbo codes) [3], where their effectiveness can be estimated by the Euclidian distance. In this paper, we propose the implementation of a fuzzy extractor based on the scheme of the so-called fuzzy commitment [2] using non-binary turbo codes. The proposed scheme has better biometric performance and implementation flexibility compared to [2,3] and has the ability to choose the type of non-binary code, arbitrary its block length and the distortion level due to data quantization to achieve the necessary confidentiality and data security.

I. SYSTEM MODEL

The proposed scheme includes two basic procedures: Enrollment and Authentication (see Fig.1). At the registration side, the m -ary Secret Password S_m enters the Non-Binary Encoder, where the encoding function $NBE(S_m) \rightarrow X_m$ add the redundant symbols for error correction,

forming framed data blocks X_m that pass through the m -ary modulator and are subtracted from a block of biometric quantized data B_q formed at the output of a Quantizer $D_m = B_q - X_m$. The quantizing interval used takes into account the power of the noise-resistant error-correcting code used and the specified level of the user data security. The obtained data block D_m is written to the Data Base and stored together with the hash $h(S_m)$ in it. At the authentication side, the subtraction $B'_q - D_m = Y_m$ for a new data block B'_q is performed, resulting in a vector Y_m , that becomes an input to the Non-Binary Decoder. The decoding function $NBD(Y_m) \rightarrow S'_m$ is applied giving the user password S'_m as the output. Next, hash function $h(S'_m)$ was compared with hash function $h(S_m)$ where compared. If they are equal, the user is successfully authenticated.

II. RESULTS AND CONCLUSIONS

In this paper we consider the use of non-binary turbo codes constructed from non-binary convolutional component codes concatenated via a random symbol interleaver mapped onto phase-shift keying (8 - PSK) constellation. The polynomials used produced coding matrix $g = [166; 176]$ over the ring $GF(8)$ for systematic 1/3-rate turbo code.

Then random secret key 8-ary S_m of length 166 was turbo encoded with terminating zeroes into resulting matrix 3×172 of X_m and then modulated into a constellation 8 - PSK. Each symbol of X_m was presented by $I - Q$ complex numbers giving framed data matrix 3×344 . To get biometrical face features the Caltech Base has been used. Data from 511 real numbers, obtained after a special mask to get the most representative components of 4464-element HOG vectors have been used as biometric raw data [2]. The biometric quantized data B_q was calculated after quantization with interval $q = 0.19635$, normalized and linearly mapped to the interval $[0, 2\pi)$ of angles presented then by two $I - Q$ components. The data block was obtained and put to the Data Base together with the hash $h(S_m)$ in it.

At the authentication stage, the subtraction $B'_q - D_m = Y_m$ for a new quantized data B'_q , was performed and a vector Y_m , was decoded after 3 iterations by the modified BCJR algorithm giving

the user password S'_m . Next, hash function $h(S'_m)$ to $h(S_m)$ where compared.

Preliminary experimental estimates of FAR and FRR resulted in values $FAR = 0\%$; $FRR \sim 0.1\%$, which is several times better than the known results for turbo codes [3].

The application of the proposed method allows to significantly improve the main performance indicators of biometric systems based on fuzzy extractors and to adjust the system parameters to

the required length of the user's secret key and the necessary level of confidentiality.

1. Dodis Y., Reyzin L., Smith A.. Fuzzy extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. EUROCRYPT, 2004. P. 523-540
2. Assanovich B. A., Veretilo Yu. N.. Biometric database based on HOG structures and BCH codes. Information Technologies and Systems 2017 (ITS-2017). Minsk, 2017. -P. 286-287
3. Maiorana E., Blasi D., Campisi P.. Biometric Template Protection Using Turbo Codes and Modulation Constellations. IEEE WIFS, 2012. P.25-30.

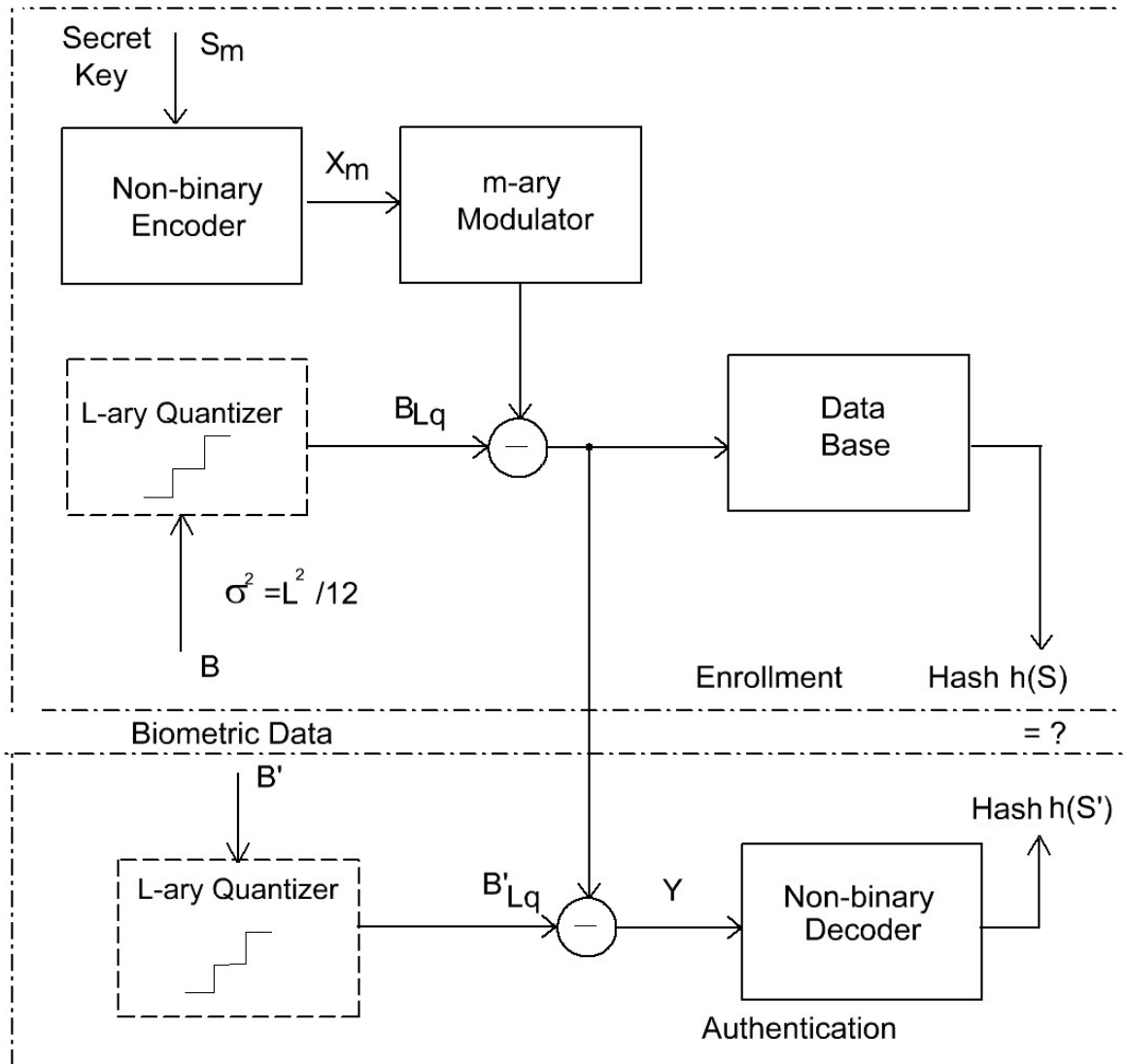


Figure 1 – System Model

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ СКЛАДИРОВАНИЯ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ АРХИТЕКТУРЫ СИСТЕМЫ СБОРА И АНАЛИЗА ДАННЫХ ПРОИЗВЕДЕНИЙ ИСТОРИЧЕСКОЙ ЦЕННОСТИ

Рудикова Л. В., Бандысик С. Ю.

Кафедра современных технологий программирования, Гродненский государственный университет имени Янки Купалы

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Гродно, Минск, Республика Беларусь

E-mail: {rudikowa, bandysik}@gmail.com

Представлен подход к построению архитектуры для системы сбора и анализа данных о произведениях исторической ценности с использованием технологий складирования данных. Предлагается общая архитектура для построения систем хранения, обработки и анализа данных, выделены основные структурные элементы системы. Лежащая в основе уровневая архитектура, позволяет построить расширяемую систему с независимыми между собой слоями. Предполагается, что представленная система будет способствовать генерации дополнительной информации о предметах исторической ценности.

ВВЕДЕНИЕ

Развитие методов записи и удешевление стоимости хранения данных способствует возрастанию объема накапливаемой информации из различных источников, которая практически никак не обработана. При более тщательном рассмотрении предметных областей, для которых используются технологии складирования данных, можно выделить средства и ресурсы, используемые для анализа накопленной информации [1].

В результате работы над объектами живописи и предметами исторических ценности [2],

многие лаборатории накапливает огромное количество данных, которые представляют собой потенциальные возможности по добыче новой, аналитической информации, используя ее можно выявить какие-либо тенденции, а также получить новую информацию, используя алгоритмы Data Mining [3-5]. Для этой цели необходима централизованная система хранения и анализа, которая будет способствовать генерации дополнительной информации о предметах исторической ценности, а также предоставлять пользователям доступ к уже накопленным данным.

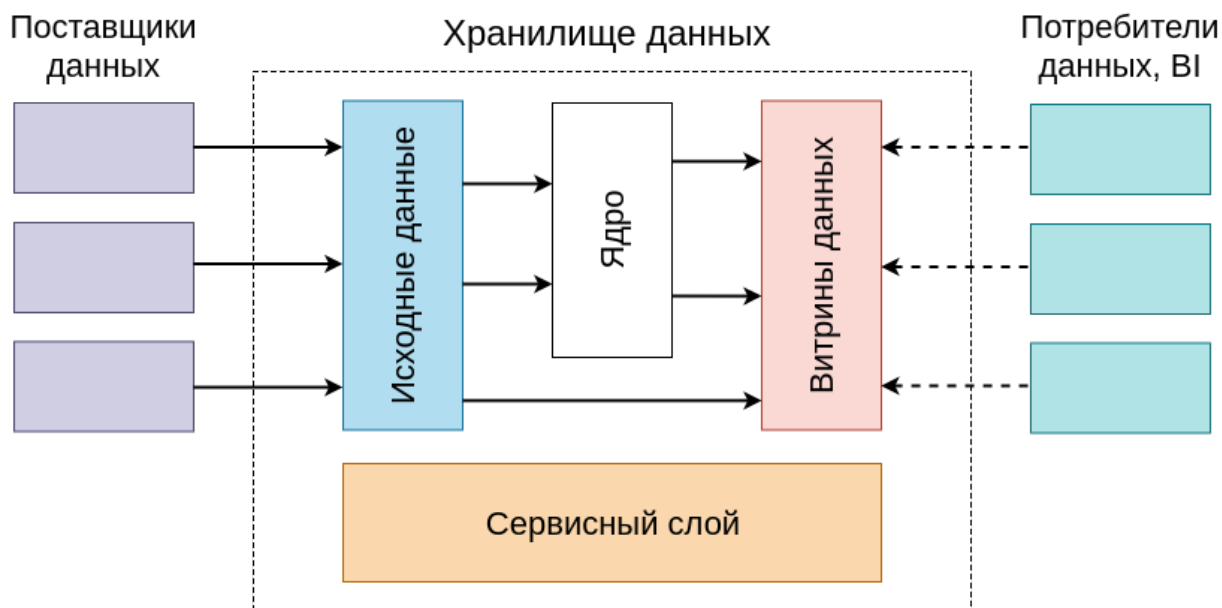


Рис. 1 – Схема архитектуры системы

ОБЩАЯ СХЕМА АРХИТЕКТУРА СИСТЕМЫ ДЛЯ СБОРА И АНАЛИЗА ДАННЫХ ПРОИЗВЕДЕНИЙ ИСТОРИЧЕСКОЙ ЦЕННОСТИ

Реализация архитектуры для систем складирования данных является непростой задачей. Для борьбы со сложностью системы необходимо использовать уровневую архитектуру, где слои реализованы независимо и абстрагированы от других. Рассматривая их по отдельности, можно более качественно выделить зоны ответственности для каждого.

Схематично можно представить следующую схему для построения системы для сбора и анализа данных произведений исторической ценности (см. рис. 1).

На схеме предложены основные слои. Слой исходных данных; Слой ядра; Слой витрин данных. Также дополнительно выделяется Сервисный слой, который отвечает за управление загрузкой в систему данных, а также за контроль качества.

Изложим некоторые характеристики основных компонентов системы.

Слой исходных данных, который отвечает за загрузку исходной информации без изменений из систем-поставщиков. Назначение слоя: позволяет данным на последующих слоях не зависеть от устройства поставщиков данных.

Слой ядра, является основным компонентом системы, который придает хранилищу данных упорядоченность, так как его главная роль – это объединение данных от нескольких поставщиков, преобразование к единым структурам. Как раз при работе на этом слое проводится важная работа по качественному обеспечению данных. К тому же здесь происходят общие преобразования, которые могут оказаться ресурсозатратными. Назначение слоя ядра: обеспечение целостности данных, а также их качества. Независимость от структуры, своими потребителями и данных поставщиков.

Слой витрин данных. Оперировать данными от надежного поставщика, используемого внутри системы, который расположен на слое ядра. Также, в некоторых случаях, возможно получение данных из слоя исходных данных. Данный способ используется в задачах, где необходим быстрый доступ к данным. Там где скорость доступа важнее качества данных. Назначение слоя: подготавливает данные, трансформируя их в удобные для анализа структуры.

Представленное логическое распределение по слоям физически может варьироваться. Для достижения быстродействия, могут быть использованы различные платформы, которые хранят, либо преобразовывают данные на слоях. Также для поддержки информационного контроля выполняющихся потоков данных, а также процессов получения информации от поставщиков, данные, представленные в таблицах для уровней

архитектуры помечаются специальными meta-атрибутами.

Сервисный слой осуществляет вспомогательные операции для всех представленных уровней, структурирует обрабатываемые потоки данных. Структуры хранения представлены областью метаданных и областью для работы с качеством данных.

ЗАКЛЮЧЕНИЕ

Приведена общая схема архитектуры для построения систем хранения, обработки и анализа данных. Также были выделены и даны краткие характеристики основных структурных элементов. Таким образом, изложенная архитектурная концепция может быть взята за основу для реализации различного вида систем, предусматривающих использование технологий хранения данных. Так как подобного рода системы позволяют собирать достаточно объемные данные из разных предметных областей, применять необходимые алгоритмы Data Mining, а также генерировать необходимые аналитические отчеты и обрабатывать данные. В частности, данную архитектуру планируется использовать для создания системы сбора и анализа данных произведений исторической ценности.

Таким образом, предлагаемое архитектурное решение представляет собой расширяемую и гибкую модульную систему в основе которой лежит сбор информации из различных источников для последующего анализа, включая генерацию новых, неочевидных зависимостей. Представленная архитектура системы предусматривает возможность расширения и гибкой модификации. На ее базе можно построить систему для использования в качестве вспомогательного инструмента для построения отчетов и проведения различного рода исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Рудикова, Л. В. Об общей архитектуре универсальной системы хранения и обработки данных практико-ориентированной направленности / Л. В. Рудикова // Системный анализ и прикладная информатика : международный научно-технический журнал.- 2017.- № 2.- С.12-19
2. Рудикова, Л. В. О разработке системы для поддержки экспертизы объектов художественной ценности / Л. В. Рудикова // Технологии информатизации и управления. Вып. 3: в 2 кн.. Кн. 2 : сборник научных статей/под. ред. А. М. Кадан, Е. А. Свицкого.- Минск: РИВШ, 2017.- С.107-117
3. Devlin, B. A. An Architecture for a Business and Information System / B. A. Devlin, P. T. Murphy. – IBM Systems Journal, 1988. – Vol 17, No 1. – P. 60–80.
4. Inmon, W. H. Building the Data Warehouse / W. H. Inmon // Third Edition. – John Wiley and Sons, Inc. New York, 2002. – 428 p.
5. Kimbell, R. The Data Warehouse Toolkit: The Complete Guide to Dimensional Data Warehouses / R. Kimbell, M. Ross // Second Edition. – J. Willey and Sons, 2002. – 447 p.

ПРИМЕНЕНИЕ ОБЛАЧНЫХ ТЕХНОЛОГИЙ ПРИ РАЗРАБОТКЕ МЕДИЦИНСКИХ ДИАГНОСТИЧЕСКИХ АППАРАТНО-ПРОГРАММНЫХ КОМПЛЕКСОВ

Боброва Т.С., Кузнецова О.В.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {t.bobrova, ovkuznetsova}@bsuir.by

В данной статье представлен обзор информации по возможностям и перспективам использования облачных технологий при разработке аппаратно-программных комплексов для проведения диагностики в области медицины и здравоохранения. Приведен пример диагностического аппаратно-программного комплекса с возможностью дистанционной обработки и анализа полученных данных.

В последнее время широкое распространение в области медицины и здравоохранения приобрели ИТ-технологии: появились электронные истории болезней, электронные очереди и электронная запись к специалистам, электронные базы данных по диспансерному учету пациентов, автоматизированная подготовка рецептов и больничных листов, консультации различного рода и т.д. Применение ИТ в системе здравоохранения привело к понятию «электронное здравоохранение» (e-health). Одной из главных тенденций в этой области выступает интеграция информационных систем в здравоохранении разного рода и назначения, причем реализованных на различных платформах, в единую систему, т.е. создание единого информационного пространства здравоохранения [1]. При этом остро встает проблема взаимодействия разнородных информационных систем между собой, в том числе диагностических, которая решается внедрением облачных технологий в медицину, поднимая ее на новый уровень.

1. ОБЛАЧНЫЕ ТЕХНОЛОГИИ В ДИАГНОСТИЧЕСКОЙ МЕДИЦИНЕ

Облачные технологии (англ. cloud computing) – модель обеспечения удобного сетевого доступа по требованию к некоторому общему фонду конфигурируемых вычислительных ресурсов, например, сетям передачи данных, серверам, устройствам хранения данных, приложениям и сервисам – как вместе, так и по отдельности, которые могут быть оперативно предоставлены и освобождены с минимальными эксплуатационными затратами или обращениями к провайдеру [2].

Применение облачных технологий в диагностической медицине важно, поскольку они помогают быстрее и точнее определить диагноз и сделать заключение, от которого зависит дальнейшее лечение пациента, становится намного проще вести и систематизировать медицинскую документацию, а также способствуют установлению информационных связей между лечеб-

ными заведениями и обмену опытом врачей-специалистов.

С развитием технологий в области диагностических медицинских приборов и оборудования, появилась необходимость обрабатывать и хранить огромные массивы различных данных, преимущественно в бумажном виде, тяжело поддающиеся обработке и анализу, что значительно снижает эффективность медицинской помощи в целом. Облачные технологии упрощают как саму процедуру диагностики различных заболеваний, требующих хранения больших объемов получаемых данных, так и оперирование ими (например, диагностика онкологических заболеваний при помощи РНК-секвенирования).

Так же незаменимо применение облачных технологий в дистанционных диагностических устройствах, позволяющих дистанционно фиксировать и транслировать в непрерывном режиме мониторинга физиологические параметры, такие как пульс, артериальное давление, уровень сахара в крови, регистрация ЭКГ-сигнала, позволяющее использовать «облачные технологии» в качестве быстрого и удобного удаленного сервера, предназначенного для сбора, хранения и обработки информации.

В таких системах данные, регистрируемые дистанционным диагностическим устройством либо датчиком, передаются на смартфон пациента с помощью беспроводной технологии Bluetooth, затем передача данных со смартфона на серверное оборудование сотового оператора происходит посредством сотовой сети с применением пакетной передачи данных. Передача данных между сервером сотового оператора и серверным оборудованием сервис-провайдера, а также устройствами диагностического центра (рабочими местами врачей) осуществляется по сети Интернет [3].

Так же возможна другая конфигурация, когда диагностирующие устройства оснащены Wi-Fi или сотовой связью, в таком случае данные с диагностирующего устройства передаются непосредственно в медицинское облако.

II. АППАРАТНО-ДИАГНОСТИЧЕСКИЕ КОМПЛЕКСЫ С ПРИМЕНЕНИЕМ ОБЛАЧНЫХ ТЕХНОЛОГИЙ

Данные, полученные в ходе диагностики пациента, должны храниться на специальных серверах, обладающих большой мощностью и емкостью. Стоимость внедрения облачных технологий в создание аппаратно-программных комплексов, обновление и обслуживание такого серверного оборудования является затратным делом, но имеет ряд положительных качеств, таких как: доступность (доступно всем пользователям), гибкость, надежность, относительная безопасность (гарантирует сохранения конфиденциальности данных, при должном использовании), большие вычислительные мощности.

К недостаткам внедрения облачных технологий в дистанционную диагностику можно отнести: необходимость постоянного соединения с сетью, специальное программное обеспечение, относительная безопасность, дороговизна оборудования.

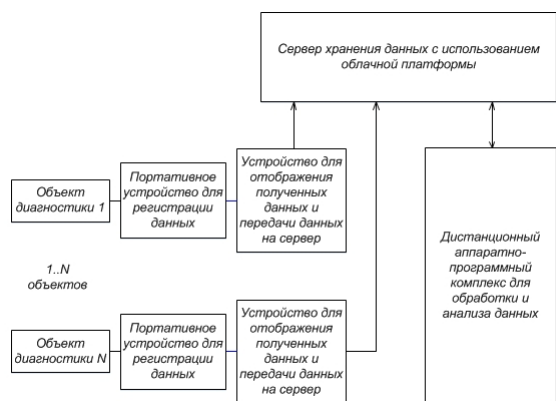


Рис. 1 – Структурная схема диагностического аппаратно-программного комплекса с применением облачных технологий

Представленный на рисунке 1 диагностический аппаратно-программный комплекс состоит из:

- портативного устройства для регистрации диагностических данных;
- устройства для отображения полученных данных и передачи данных на сервер;
- сервера хранения данных, с использованием облачной платформы;
- аппаратно-программного модуля обработки и анализа данных.

В качестве портативного устройства для регистрации диагностических данных могут выступать различного рода датчики, сенсоры, портативные измерители, медицинские приборы, предназначенные для получения биомедицинских сигналов. С портативного устройства данные по каналу Bluetooth, через USB-порт или другим каналам связи передаются на базовое устройство отображения и передачи на сервер.

В роли этого устройства могут выступать персональный компьютер, ноутбук, планшет и даже смартфон, имеющие доступ к сети Internet для выгрузки данных в хранилище на сервере. Здесь же может проводиться предварительная обработка данных, добавление данных о пациенте и описание полученных данных. Сервер хранения данных организован с использованием облачной платформы, поддерживающей необходимый функционал для работы с полученными данными. В качестве облачной платформы возможно Google Cloud Platform, Amazon Web Services, Microsoft Azure и другие. На сервер могут одновременно поступать данные от нескольких объектов диагностики, расположенных в различных местах и учреждениях здравоохранения, что значительно ускоряет процесс сбора, обработки и анализа данных, а также делает их хранение более удобным и безопасным. Для дальнейшей обработки и анализа данных используется аппаратно-программный модуль обработки и анализа данных, входящий в состав комплекса, который представляет собой вычислительное устройство с установленным на него необходимым программным обеспечением и может быть реализован на базе персонального компьютера (ноутбука или планшета). Все устройства имеющие доступ к данным могут успешно проводить их обработку, так же на сервер можно передавать уже обработанные данные и результаты их анализа, для использования другими пользователями. Технология с применением облачной платформы позволит повысить точность результатов, получаемых в ходе диагностики, структурированно хранить, эффективно обрабатывать и анализировать большие объемы данных.

ЗАКЛЮЧЕНИЕ

Использование облачных технологий при разработке диагностических аппаратно-программных комплексов и систем позволяет работать с большими объемами данных, проводить одновременно диагностику нескольких пациентов, проводить анализ и обработку полученных данных дистанционно, обеспечивает безопасное структурированное хранение данных, чем значительно облегчает поиск нужной информации по сложным запросам и упрощает работу врачей.

1. Каменщиков А.А. Облачные технологии и interoperability информационных систем в здравоохранении // Журнал радиоэлектроники: электронный журнал. 2013. - № 2 [Электронный ресурс]. URL: http://j.re.cplire.ru/j_re/feb_13/11/text.pdf.
2. Mell, Peter and Grance, Timothy. The NIST Definition of Cloud Computing (англ.). Recommendations of the National Institute of Standards and Technology. NIST (20 October 2011)
3. Кузнецов В.И., Тараканов С.А., Рыжаков Н.И., Рассадина А.А. Применение облачных технологий в медицинских дистанционных диагностических устройствах // Врач и информационные технологии. 2012. - N 5.-С.68-72. Библ. 17 назв.

МЕТОДИКА ПЕРЕКРЕСТНОЙ ОЦЕНКИ УЯЗВИМОСТЕЙ И УГРОЗ БЕЗОПАСНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА

Буй П. М., Кульгавик С. Г.

Кафедра «Автоматика, телемеханика и связь», Белорусский государственный университет транспорта
Гомель, Республика Беларусь

E-mail: pashabuoy@rambler.ru, kalashnikovn27.sk@gmail.com

В статье обоснована необходимость обеспечения как информационной, так и функциональной безопасности информационных систем железнодорожного транспорта, предложена методика перекрестной оценки угроз и уязвимостей, приведены необходимые критерии и система выставления баллов экспертами.

ВВЕДЕНИЕ

Для Республики Беларусь железнодорожный комплекс имеет особое стратегическое значение, являясь связующим звеном единой экономической системы и обеспечивая стабильную деятельность промышленных предприятий. Кроме того, это еще и самый доступный вид транспорта для граждан республики. Все это способствует тому, что Белорусская железная дорога обязана обеспечить потребности государства, юридических и физических лиц в железнодорожных перевозках, а также работах и услугах, оказываемых железнодорожным транспортом. В рамках стремительной информатизации и компьютеризации общества Белорусская железная дорога не в состоянии качественно выполнять поставленные перед ней задачи, не прогрессируя вместе с обществом. Внедрение передовых и вместе с тем надежных технологий по ее информатизации является одной из первостепенных задач. Вместе с тем процессы информатизации и компьютеризации, а также использование современных сетевых технологий при организации управления на Белорусской железной дороге таят в себе множество потенциальных опасностей, область реализации которых касается исключительно сферы высоких технологий. При отсутствии адекватной системы защиты опасности такого рода могут привести к нарушению штатной работы систем управления и, как следствие, ухудшению уровня безопасности грузо- и пассажироперевозок. В таких условиях обязательным является проведение анализа этих опасностей характерных как для самих информационных систем, так и для среды их функционирования. В сфере железнодорожного транспорта довольно часто информационные системы используются не только для обработки информации, но и для организации автоматизированных систем управления технологическими процессами (АСУ ТП), в некоторых из которых может вообще отсутствовать информация, предоставление и/или распространение которой ограничено. Безопасность таких информационных систем – это их защищенность от случайного или преднамеренного вме-

шательства в штатный процесс их функционирования. В общем случае речь идет о функциональной безопасности информационной системы, когда важным является выполнение системой поставленных перед ней задач. Если же эти задачи связаны с хранением и обработкой информации, предоставление и/или распространение которой ограничено, то в таком случае речь идет об информационной безопасности. Для систем управления на железнодорожном транспорте именно нарушение функциональной безопасности становится более опасным.

I. КРИТЕРИИ ОЦЕНКИ

Для адекватной оценки уровня безопасности информационной системы важно проводить анализ ее уязвимостей и потенциальных угроз ее безопасности. При этом следует охватить существенное их количество. Такой подход дополнительно усложняется, если принять во внимание, что несколько угроз могут быть реализованы через одну и ту же уязвимость и аналогично несколько уязвимостей могут стать причиной реализации одной и той же угрозы. В таких обстоятельствах целесообразно использовать методику перекрестной оценки угроз безопасности информационных систем и их уязвимостей [1]. Это позволит учесть очевидную взаимосвязь угроз и уязвимостей, являющуюся обязательным условием реализации любой угрозы, а также вопросы не только информационной, но и функциональной безопасности, которые зачастую остаются в «тени» при использовании существующих методов, ориентированных на оценку исключительно информационной безопасности. Методика перекрестной оценки угроз безопасности информационных систем и их уязвимостей опирается на методику экспертных оценок. В связи с этим квалифицированные эксперты должны определить и выставить баллы следующим специальным критериям для каждой пары «угроза-уязвимость» дискретно в диапазоне от 1 до 10 [2]:

- критерий C_1 (от англ. Criterion) – возможность возникновения источника угрозы в достаточном окружении от информацион-

- ной системы для реализации угрозы через уязвимость;
- критерий C_2 – степень готовности источника угрозы воспользоваться уязвимостью информационной системы и реализовать угрозу;
 - критерий C_3 – распространенность уязвимости по информационной системе или частота ее появления;
 - критерий C_4 – доступность уязвимости для реализации угрозы ее источником;
 - критерий C_5 – фатальность от реализации угрозы источником угрозы через уязвимость информационной системы.

Принцип выставления баллов для первых четырех критериев прост: чем в большей степени появляется критерий, тем большего балла он заслуживает. Для учета вопросов как информационной, так и функциональной безопасности для пятого критерия рекомендуются представленные на рисунке 1 значения баллов и соответствующие им уровни нарушения безопасности информационных систем исходя из соображений первостепенной важности обеспечения функциональной безопасности.

II. МЕТОДИКА ПЕРЕКРЕСТНОЙ ОЦЕНКИ

При проведении перекрестной оценки уязвимостей и угроз безопасности информационной системы необходимо:

1. Определить совокупности угроз и уязвимостей безопасности информационной системы;
2. Увязать между собой угрозы и уязвимости, установив потенциальную реализацию первых через вторые;

3. Перевести в резерв несвязанные уязвимости и угрозы;
4. Вычислить коэффициенты опасности реализации каждой угрозы через каждую связанную с ней уязвимость (формула расчета коэффициента опасности приведена в источнике [1]);
5. Для каждой угрозы определить коэффициент опасности с учетом возможной ее реализации через некоторые уязвимости из перечня уязвимостей, определенного в пункте 1 (формула расчета коэффициента опасности угрозы приведена в источнике [2]);
6. Для каждой уязвимости определить коэффициент опасности с учетом возможной реализации через нее некоторых угроз из перечня угроз, определенного в пункте 1 (формула расчета коэффициента опасности уязвимости также приведена в источнике [2]);
7. Произвести ранжирование угроз и уязвимостей, определив тем самым наиболее опасные из них.

1. Буй, П. М. Методика перекрестной оценки угроз безопасности информационных систем и их уязвимостей / П. М. Буй, С. Г. Кульгавик // Комплексная защита информации: материалы XXIII Международной научно-практической конференции, Суздаль, 22-24 мая 2018г. – Суздаль: НИИ ТЗИ, 2018.
2. Буй, П. М. Методика перекрестной оценки угроз и уязвимостей безопасности объектов информатизации железнодорожного транспорта / П. М. Буй, С. Г. Кульгавик // Вестник БелГУТа: Наука и транспорт – 2017. – №2 (35). – С. 40 – 43.

Балл, выставаемый экспертом	Уровни нарушения безопасности информационных систем				
	нарушение доступности информации	нарушение конфиденциальности информации	нарушение целостности информации	частичное нарушение функциональной безопасности	выход из строя информационной системы
1	+				
2		+			
3		+	+		
	+	+	+		
4	+	+	+		
5				+	
6	+			+	
7		+		+	
			+	+	
8		+	+	+	
	+	+		+	
	+		+	+	
9	+	+	+	+	
10					+

Рис. 1 – Значения баллов критерия фатальности реализации угрозы через уязвимость

К ВОПРОСУ АВТОМАТИЗАЦИИ НАСТРОЙКИ АЛАР КПАМ: ФИКСАЦИЯ ИЗМЕНЕНИЯ СОПРОТИВЛЕНИЯ

Бухаров Д. С.

Служба автоматизированных систем диспетчерского управления, филиал АО «СО ЕЭС» «Региональное диспетчерское управление энергосистемы Иркутской области»

Иркутск, Российская Федерация

E-mail: bukharovds@gmail.com

Реализованы методы определения принадлежности точек годографа вектора сопротивления к характеристике срабатывания устройств АЛАР, расположения органа направления мощности, вычисления формы характеристики с учетом условий чувствительности устройства АЛАР. Программное обеспечение позволяет оценить время нахождения годографа в характеристике срабатывания и графически отобразить результаты расчетов для возможности анализа полученных результатов.

ВВЕДЕНИЕ

Настройка устройств автоматической ликвидации асинхронного режима (АЛАР) – трудоемкий процесс, сопряженный со значительным количеством ручных операций, автоматизация выполнения которых имеет немаловажное значение. Получение адекватной настройки устройств АЛАР напрямую связано с многократной имитацией работы устройств АЛАР на совокупности исходных данных, характеризующих собой переходные процессы с возникновением асинхронных режимов (АР).

1. МОДЕЛИРОВАНИЕ РАБОТЫ АЛАР

Моделирование работы устройств АЛАР и автоматическое вычисление параметров их характеристики срабатывания – ключевые процессы, заложенные в основу разрабатываемого программного обеспечения (ПО).

Исходными данными при настройке характеристики являются массивы значений времени, активного R и реактивного X сопротивления, полученных в результате расчета переходного процесса. Каждый массив содержит в себе по несколько тысяч временных отсечек и парных им значений R и X . При этом различие в мощности данных массивов может составлять до нескольких тысяч значений.

В основу разрабатываемого ПО автоматической настройки устройств АЛАР положена итеративная процедура вычисления характеристики, которая разработана с учетом рекомендаций разработчика устройств КПАМ [1].

Срабатывание устройства АЛАР фиксируется при последовательном прохождении годографа АР через чувствительный (ЧО) и грубый орган (ГО) характеристики и пересечении органа направления мощности (ОНМ) (рис. 1), при этом вычисляется время прохождения годографа через область ЧО: $T = t_2 - t_1$, где t_1 и t_2 – моменты времени, в которые зафиксированы соответственно первая и последняя точка последовательности, принадлежащей области ЧО.

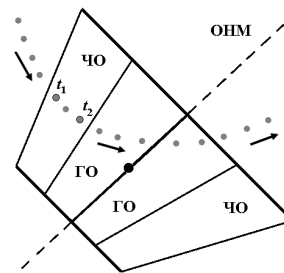


Рис. 1 – Моделирование работы устройства АЛАР

Если $T \geq T_{act}$ (T_{act} – минимально необходимое время для фиксации АР), то отмечается работа автоматики и годограф маркируется «пойманным».

II. ОПРЕДЕЛЕНИЕ ПРИНАДЛЕЖНОСТИ ТОЧКИ К ЗАДАННОЙ ГЕОМЕТРИЧЕСКОЙ ФИГУРЕ

Поскольку характеристика устройства АЛАР имеет трапециевидную форму, то она без особых проблем представляется в виде двух треугольников.

Исследованы алгоритмы определения принадлежности точки к заданному треугольнику, основанные на методе сравнения площадей, методе относительности координат, векторном методе и методе трассировки лучом.

В ходе вычислительного эксперимента наилучший результат показал метод относительности координат.

Определение принадлежности точки к заданной геометрической фигуре – одна из самых трудоемких и затратных по вычислительным ресурсам процедура. Поэтому применение этого метода в чистом виде является неэффективным.

Существенно меньшее количество вычислительных операций выполняется при определении принадлежности точки к прямоугольнику. Выполнено комбинирование двух методов с целью снижения вычислительных затрат.

Суть комбинирования следующая: трапециевидная характеристика описывается прямоугольником; в первую очередь для любой точки выполняется проверка принадлежности к этому

описывающему прямоугольнику; если некоторая точка оказывается внутри описывающего прямоугольника, то принадлежность определяется по методу относительности координат.

Количество точек, лежащих внутри описывающего прямоугольника, составляет от 10% до 20% от общего числа. Время вычисления сокращается приблизительно на 70%.

III. ЭТАПЫ НАСТРОЙКИ АЛАР

- Определение положения ОНМ: осуществляется методом деления пополам, который специально модифицирован для решения текущей задачи. Метод позволяет за конечное число операций вычислить наилучший угол наклона ОНМ. Угол наклона ОНМ оказывает существенное влияние на конечную форму характеристики.
- Определение длин оснований трапеции ГО и ЧО: осуществляется шаговым методом с инерционной составляющей, позволяющей выходить из области притяжения локальных минимумов и получать такие длины оснований, при которых охватывается наибольшее количество АР. Бесконечное увеличение длин оснований невозможно, поскольку величины R и X изменяются в конечном диапазоне, охват которого хотя бы на половину приводит к невозможности фиксации АР.
- Корректировка высоты и длин боковин трапеции по условию чувствительности: осуществляется наращиванием длин на заданную величину согласно рекомендациям разработчика устройств АЛАР [1]. Условие чувствительности для боковины/высоты соблюдается, если точка пересечения годографа АР и боковины/высоты делит по-

следних на отрезки более 0,1 от длины этой боковины/высоты.

Для получения оптимальной настройки устройства АЛАР необходимо многократное моделирование работы устройства АЛАР после каждого изменения характеристики. Это необходимо для вычисления таких параметров, которые обеспечивали бы корректную работу самого устройства АЛАР в условиях промышленной эксплуатации на объектах электроэнергетики.

IV. ЗАКЛЮЧЕНИЕ

Все алгоритмы реализованы в виде ПО «Расчет параметров устройства АЛАР» (рис. 2). Вычисление характеристики устройства АЛАР при обработке 500 годографов на компьютере с четырехядерным процессором Intel(R) Core(TM) i7-5600 CPU 2.60 GHz под управлением операционной системы Windows 7 выполняется за 10 минут. За весь цикл автоматической настройки характеристики выполняется от 10^7 до 10^8 проверок принадлежности точек годографов АР к характеристике.

Снятие со специалиста вычислительной рутинной и предоставление возможности рассмотрения существенно большего количества АР, чем при «ручной» настройке, позволяет перевести процесс принятия решения по выбору характеристики на качественно иной уровень: исключается необходимость экспертной оценки исходных данных с целью получения АР с наибольшим влиянием на энергосистему; появляется возможность исследования большого числа АР, исключая человеческий фактор при выполнении вычислений.

1. Комплекс противоаварийной автоматики многофункциональный. 656455.206-01 РЭ. Новосибирск: Институт автоматизации энергетических систем, 2013. – 60 с.

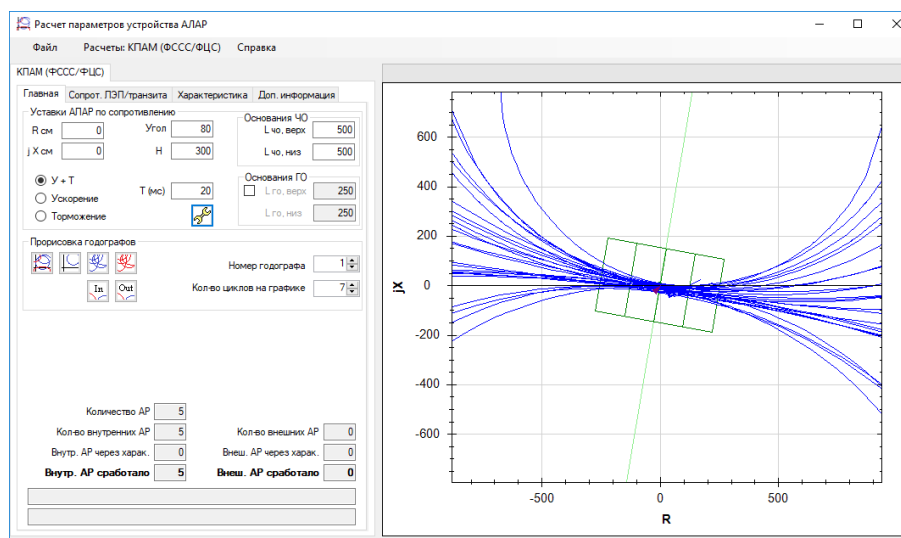


Рис. 2 – Интерфейс ПО «Расчет параметров устройства АЛАР»

К ВОПРОСУ АВТОМАТИЗАЦИИ НАСТРОЙКИ АЛАР КПАМ: ФИКСАЦИЯ КОЛЕБАНИЯ ТОКА

Бухаров Д. С.

Служба автоматизированных систем диспетчерского управления, филиал АО «СО ЕЭС» «Региональное диспетчерское управление энергосистемы Иркутской области»

Иркутск, Российская Федерация

E-mail: bukharovds@gmail.com

Представлены два метода поиска экстремумов в исходном массиве данных, характеризующем собой график колебания тока. Исходные данные обладают особенностью (локальные колебания малой амплитуды), препятствующей применению для решения задачи вертикального и горизонтального сканирования. Адекватное решение задачи дает метод равномерного поиска (метод сканирования), но при этом снижается точность решения из-за потери большого количества исходных данных. Автором предложен метод решения задачи без потери данных. Метод основан на разделении исходного множества данных на пересекающиеся подмножества и анализе всех данных на малых интервалах времени.

ВВЕДЕНИЕ

В задаче автоматической настройки устройств АЛАР, выполненных по принципу фиксации колебания тока [1], можно выделить три этапа: вычисление экстремальных значений по каждому исследуемому графику колебания тока; вычисление уставок тока срабатывания I_{cp} и возврата I_{voz} ; вычисление наибольшего периода колебания T_{max} .

Результат решения каждого этапа является входными данными для следующего: для определения наибольшего периода колебания тока T_{max} необходимы значения токов I_{cp} и I_{voz} ; для вычисления токов I_{cp} и I_{voz} необходимо знать все максимумы и минимумы для каждого графика переходного процесса (рис. 1).

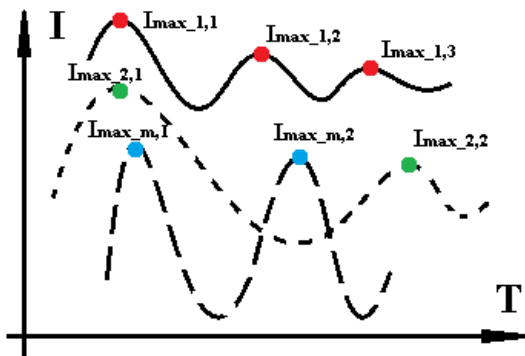


Рис. 1 – Искомые точки максимума

Наибольший интерес представляет вычисление экстремальных значений по каждому исследуемому графику колебания тока. Для поиска экстремумов могут применяться различные методы поиска по заданной маске, вертикального и горизонтального сканирования графика, стохастические методы, методы с предварительным преобразованием исходных данных. Однако применение большинства этих методов может оказаться малоэффективным по причине наличия у графиков тока «локальных» колебаний

(рис. 2), способных существенно исказить результат настройки устройства АЛАР.

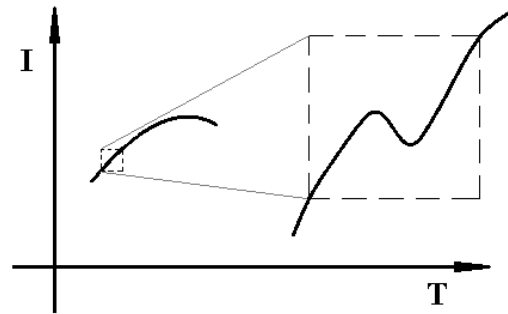


Рис. 2 – Локальные колебания графика

1. МЕТОДЫ ПОИСКА ЭКСТРЕМУМА

Наиболее простым в реализации методом поиска экстремумов является метод равномерного поиска. Суть метода заключается в преобразовании исходного графика (рис. 3,а) в новый график (рис. 3,б) с равномерным шагом по времени Δt . Шаг Δt принимается достаточно большим, что позволяет избежать влияния «локальных» колебаний на конечный результат, однако такой подход ведет к потере части данных.

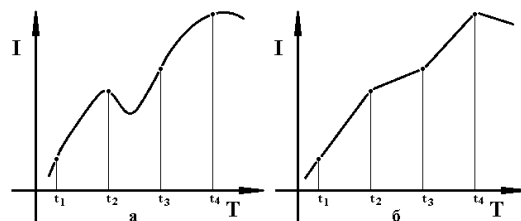


Рис. 3 – Равномерный поиск

Решить задачу с сохранением всех исходных данных можно применив метод с разбиением исходных данных на пересекающиеся подмножества (рис. 4). Суть метода заключается в разбиении всех данных $I = \{I(t_j) : j = \overline{1, n}\}$ на подмножества $sub_c = \{I_c(t_{c,j}) : j = \overline{1, n_c}\}$, где $I(t_j)$ – значение тока в момент времени t_j ,

n – общее количество замеров, $c = \overline{1, v}$ – номер подмножества, v – количество подмножеств, n_c – количество точек в c -ом подмножестве.

Разбиение выполняется на равные подмножества sub_c так, чтобы 50% элементов каждого подмножества i содержалось в подмножестве $i + 1$. Такое дублирование данных необходимо для анализа динамики изменения тока на каждом интервале sub_c и соотнесения выявленных экстремумов друг с другом.

Так например найденный максимум в подмножестве sub_i может являться минимумом в подмножестве sub_{i+1} , что позволяет сделать вывод о необходимости игнорировать данное решение и не включать в искомое множество экстремумов.

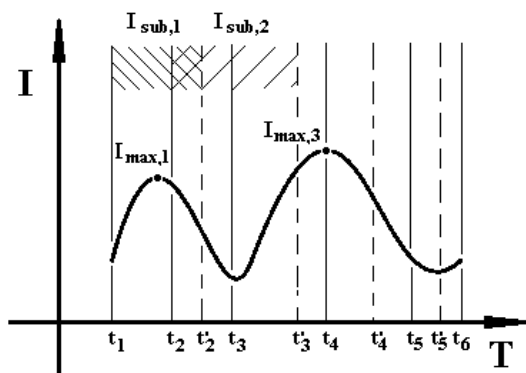


Рис. 4 – Разбиение на подмножества

Также для поиска экстремумов возможно применение методов вертикального и горизонтального сканирования, однако при этом потребуются исключение из исходных данных «локальных» колебаний (рис. 2).

Вычисленные множества минимумов $I_{min,m}$ и максимумов $I_{max,m}$ для всех исследуемых m

графиков колебания тока используются для решения задач:

$$I_{minmax} = \min(I_{max,i}),$$

$$I_{maxmin} = \max(I_{min,i}),$$

$$i = \overline{1, m}.$$

Пример вычисления данных величин в ПО «Расчет параметров устройства АЛАР» приведен на рис. 5.

II. ЗАКЛЮЧЕНИЕ

Наилучшей скоростью решения задачи обладает метод равномерного поиска: время решения задачи на 27% меньше, чем у метода с разбиением на подмножества. Однако применение метода равномерного поиска сопряжено с потерей существенной части данных, что влечет за собой снижение точности решения.

Для решения рассмотренной задачи с учетом всех исходных данных целесообразно использовать метод с разбиением на пересекающиеся подмножества, позволяющий определить точки минимума и максимума без потери точности. Данный метод позволяет исключить из конечного решения локальные экстремумы (проблема локальных колебаний), которые способны существенно исказить конечную настройку устройства АЛАР, выполненных по принципу фиксации колебания тока.

1. Комплекс противоаварийной автоматики многофункциональный. 656455.206-01 РЭ. Новосибирск: Институт автоматизации энергетических систем, 2013. – 60 с.
2. Гоник, Я. Е., Иглицкий, Е. С. Автоматика ликвидации асинхронного режима / Я. Е. Гоник, Е. С. Иглицкий // М.: Энергоатомиздат, 1984. – 112 с

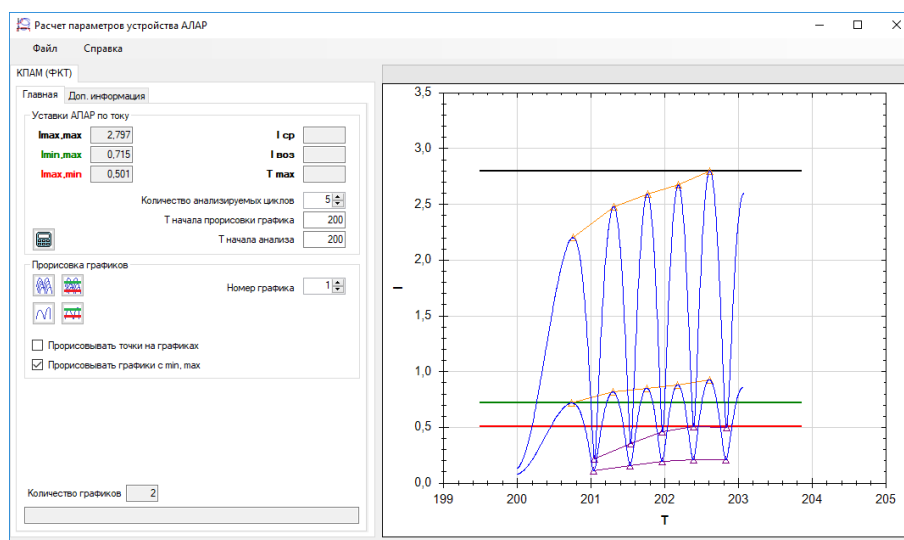


Рис. 5 – Интерфейс ПО «Расчет параметров устройства АЛАР»

КЛАССИФИКАЦИЯ ВАКАНСИЙ С ЦЕЛЬЮ ПОСЛЕДУЮЩЕЙ ОПТИМИЗАЦИИ ПУБЛИКАЦИИ ОБЪЯВЛЕНИЙ

Быстрова М. В., Козадаев И. А.

Факультет радиофизики и компьютерных технологий Белорусского государственного университета

Минск, Республика Беларусь

E-mail: mvb1610@gmail.com, kozadaeff@mail.ru

В работе представлен и проанализирован подход к классификации данных, базирующийся на методе обработки текстов на естественном языке с помощью инструментов ИАТ. Для непосредственной классификации был выбран предварительно размеченный набор данных, содержащий текстовые описания вакансий. Была проанализирована устойчивость и точность данного подхода.

ВВЕДЕНИЕ

Согласно статистическим данным, заполнение большинства рабочих мест происходит с помощью публикаций объявлений о вакансиях. Успех такого набора зависит от того, как компания преуспела в составлении соответствующего объявления. Важно знать критерии, которые способны заинтересовать потенциального сотрудника, или же, другими словами, ценность предложения для работника.

В данной работе проблема классификации вакансий рассмотрена как задача интеллектуального анализа данных, для решения которой были предложены наиболее подходящие методы предварительной обработки текста публикаций, а также выбран оптимальный алгоритм определения близости публикаций в векторном пространстве [1].

I. ОСНОВНЫЕ ЭТАПЫ ПОСТРОЕНИЯ КЛАССИФИЦИРУЮЩЕЙ МОДЕЛИ

В ходе построения классифицирующей модели были выполнены три нижеприведённых этапа (см. рис. 1).

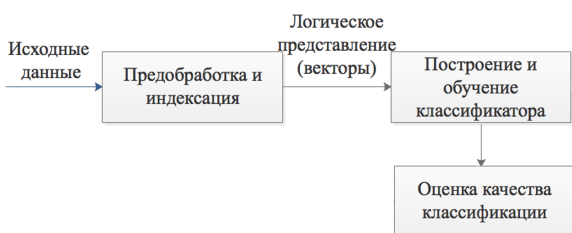


Рис. 1 – Этапы построения классифицирующей модели

II. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА И ИНДЕКСАЦИЯ

Опубликованные вакансии представляют собой небольшие текстовые документы различной длины. Для того, чтобы использовать эту информацию в качестве обучающей выборки, была произведена *токенизация*, то есть выделение в тексте слов, чисел, границ предложений

и иных *токенов* (или *термов*), а также были удалены семантически нейтральные слова такие, как предлоги и союзы. После этого к полученной информации нами была добавлена служебная метаинформация, которая содержит обучающие метки, необходимые для работы классификатора.

Текст документа представляется в виде мультимножества термов [2] и обозначается $d_j \in D$, где D - множество всех документов, присутствующих в выборке. Множество всех термов $T = \{\tau_0, \tau_1, \dots, \tau_{|T|}\}$. Каждому терму $\tau_i \in T$ ставится в соответствие некоторый вес w_{ij} , характеризующий встречаемость данного терма в тексте $d_j \in D$. Логическое представление принято обозначать вектором $\vec{d}_j = \{w_{0j}, w_{1j}, \dots, w_{|T|j}\}$, где каждый w_{ij} - вес τ_i терма в документе \vec{d}_j .

В итоге было получено n -мерное пространство векторов, которое принято называть пространством признаков для класса данных D . Таким образом каждый документ является точкой в пространстве признаков.

$$IDF(\tau_i, D) = \left(\frac{|D|}{|d_i \supset \tau_i|} \right), 0 \leq i \leq |T|, \quad (1)$$

где $|D|$ - количество документов в классе, $|d_i \supset \tau_i|$ - количество документов, в которых встречается терм τ_i .

После этого были применены методы уменьшения размерности термов для обеспечения приемлемого времени работы алгоритма. В рамках этой работы были установлены следующие правила, при выполнении которых терм считается неинформативным:

1. Встречаемость терма в выборке меньше некоторого числа n ;
2. Терм имеет большое математическое ожидание M_{f_i} и маленькую дисперсию D_{f_i} . Конкретные значения порогов задаются исходя из конкретных условий;
3. Имеет маленький информативный вес.

III. ПОСТРОЕНИЕ И ОБУЧЕНИЕ КЛАССИФИКАТОРА

После формирования и предварительной обработки тренировочного набора документов следуют выбор и построение классифицирующей модели, архитектурными компонентами которой были выбраны методы Distributed Memory (распределенная память, DM) и Distributed Bag of Words (распределенный мешок слов, DBOW).

- DM прогнозирует слово по известным предшествующим словам и вектору абзаца;
- DBOW прогнозирует случайные группы слов в абзаце на основании вектора абзаца.

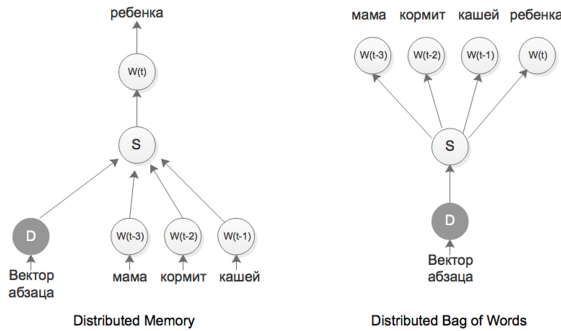


Рис. 2 – Схемы методов Distributed Memory и Distributed Bag of Words

Принцип работы этих методов заключается в нахождении связей между контекстами слов согласно предположению, что слова, находящиеся в похожих контекстах, являются семантически близкими. Формально задача представляет собой максимизацию косинусной близости между векторами слов, которые появляются в близких контекстах, и минимизация косинусной близости между векторами слов, которые не появляются рядом друг с другом: $\min\left\{\frac{w_v \times w_c}{\sum w_{cl} \times w_v}\right\}$, где w – слова контекста, w_v – целевое слово, w_{cl} – другие контексты. $w_v \times w$ – близость слов контекста и целевого слова, $w_{cl} \times w_v$ – близость всех других контекстов и целевого слова.

Решить задачу такой минимизации (маскизации) можно с помощью обычной нейронной сети прямого распространения, требующей, чтобы входные векторы были фиксированной длины. Однако, если векторные представления предложений образовывать за счет склеивания соответствующих представлений слов, на выходе всегда будут получаться векторы разной длины.

В качестве решения этой проблемы был выбран свёрточный фильтр, идея которого заключается в том, что каждому нейрону подается на вход два (или более) слова, причем для каждого последующего нейрона вход сдвигается на одно

слово. Например, первому нейрону на вход подается слово 1 и слово 2, второму – слово 2 и слово 3, и т.д. На выходе имеется предложение, которое в два или в N (количество нейронов входного слоя) раз короче исходного.

IV. ОЦЕНКА КАЧЕСТВА РАБОТЫ КЛАССИФИКАТОРА И АНАЛИЗ РЕЗУЛЬТАТОВ

В итоге была спроектирована и реализована система, которая позволяет классифицировать вакансии с использованием рассмотренных выше подходов. Для оценки качества работы модели была использована F -мера:

$$F(u) = \frac{2 * p(u) * r(u)}{p(u) + r(u)}, \quad (2)$$

где $r(u) = \frac{|u \cap v|}{|v|}$ – полнота (recall) классификации по классу, то есть отношение количества документов, для которых классификатор правильно определил класс, к общему количеству документов класса, определенному без классификатора; $p(u) = \frac{|u \cap v|}{|u|}$ – точность (precision), показывающая отношение количества документов, для которых классификатор правильно определил класс, к количеству документов, которые классификатор отнес к данному классу.

V. ЗАКЛЮЧЕНИЕ

Рассмотренные подходы были использованы для проектирования реальной системы классификации вакансий. Python использовался в качестве языка реализации. Методы DM и DBOW при сравнительно низкой вычислительной сложности позволили получить среднюю точность порядка 70%.

Несмотря на то, что нами были получены достаточно хорошие результаты, зачастую при использовании более сложных текстов на естественном языке современные технологии ИАТ не обладают высокими устойчивостью. Это связано с тем, что на текущий момент ИАТ не могут в точности определять семантику текста, а также анализировать сложные зависимости между разными его частями.

1. Быстрова, М. В. Классификация вакансий с целью последующей оптимизации публикации объявлений // Научное сообщество студентов XXI столетия: Технические науки: сб. ст. по мат. LXIV междунар. студ. науч.-практ. конф. 18 июня 2018. № 4(63)
2. Агеев, М. С. Методы автоматической рубрикации текстов, основанные на машинном обучении знаниях экспертов / М. С. Агеев // Либроком (Editorial URSS), 2004. – 106 с.

ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ ДИФФЕРЕНЦИАЛЬНО ВЫРАЖЕННЫХ ГЕНОВ И ОБОГАЩЕННЫХ БИОЛОГИЧЕСКИХ ФУНКЦИЙ РАКОВЫХ КЛЕТОК

Чепелева М. К., Яцков Н. Н., Назаров П. В.

Кафедра системного анализа и компьютерного моделирования, Белорусский государственный университет
Luxembourg Institute of Health

Минск, Республика Беларусь; Luxembourg, Luxembourg

E-mail: maryna.chepeleva@gmail.com, yatskou@bsu.by, petr.nazarov@lih.lu

В работе исследуется статистическая значимость результатов анализа данных секвенирования молекул РНК плоскоклеточного рака лёгкого, полученных в разных научно-исследовательских лабораториях. Рассматриваются алгоритмы поиска дифференциально выраженных генов и обогащенных биологических функций, оценивается статистическая устойчивость групп генов и биологических функций к изменению объема экспериментальной выборки.

ВВЕДЕНИЕ

Рак – заболевание, характеризующее неконтролируемым ростом клеток, возникающее под влиянием онкогенных факторов, реализующих свое действие через генетический аппарат клетки. Секвенирование РНК – технология определения первичной структуры молекулы РНК, позволяющая получить количественную меру экспрессии генов и предоставляющая экспериментальные данные для интеллектуального анализа. Гены, уровень экспрессии которых оказывается разным у тестируемых и контрольных образцов, называются дифференциально выраженными и подлежат дальнейшим исследованиям [1].

Биологические функции представляют собой комплексный биологический феномен, обусловленный набором генов. Биофункция формирует иерархическую структуру с другими и включает в свою аннотацию набор генов, определяющих ее работу [2]. Биофункция обогащена дифференциально выраженными генами, если список её генов содержит значительно больше дифференциально выраженных генов, чем предсказано исходя из гипергеометрического распределения. Зная список обогащенных функций, можно точно определить, в каких процессах происходят изменения при заболевании человека.

Секвенирование РНК опухолевых тканей проводят во многих научных лабораториях мира. Часто результаты анализа достаточно сильно отличаются и не подтверждают ранее полученные выводы. Поэтому важной задачей является оценка адекватности результатов анализа наборов экспериментальных данных, полученных из разных научных лабораторий. Цель данного исследования – оценка устойчивости групп генов и биологических функций к изменению объема выборки пациентов. Для этого был проведён анализ экспериментальных данных геномного секвенирования, полученных из различных научных лабораторий, с использованием алгоритмов поиска

дифференциально выраженных генов и биологических функций, обогащенных данным генами.

I. АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ

В исследовании использовались данные экспрессии генов плоскоклеточного рака легкого TCGA LUSC-dataset [3]. После фильтрации и нормализации к данным применялся алгоритм поиска дифференциально выраженных генов на основе модифицированного критерия Стьюдента. В алгоритме производится построение матрицы, которая хранит информацию о типе каждого образца. Для каждого гена строится линейная модель:

$$E[y_j] = X\alpha_j,$$

где y_j – данные экспрессии для гена j , X – расчетная матрица, α_j – вектор коэффициентов или параметров линейной регрессионной модели. Коэффициенты пересчитываются два раза с учетом типа образца и применением эмпирического метода Байеса [4]. Осуществляется проверка гипотез о дифференциальной выраженности генов с помощью критерия Стьюдента для каждого гена и для каждого образца, корректируются значения. Результатом анализа является список дифференциально выраженных генов.

Вторая часть анализа заключается в поиске биологических функций, обогащенных найденными дифференциально выраженными генами. Анализ выполнялся с использованием алгоритма на основе точного критерия Фишера [5]. Входными данными являются список дифференциально выраженных генов с рассчитанными на предыдущем шаге статистиками, а также база биологических функций. Вычисляется точный критерий Фишера с учетом иерархических отношений биофункций, определяется степень «обогащения» биофункций.

II. ИССЛЕДОВАНИЕ УСТОЙЧИВОСТИ

Для сравнения результатов анализа наборов данных разных размеров используются две

метрики сходства: индекс Жаккара

$$J = \frac{(A \cap B)}{(A \cup B)}$$

и косинусное расстояние – косинус угла между двумя многомерными векторами весов в пространстве биологических функций [6]. Косинусное сходство позволяет помимо пересечения учитывать степень обогащения при сравнении двух списков биологических функций.

Для оценки значимости одного пациента в малой выборке методом перекрестной проверки проводился анализ возможных вариантов выбора девяти из десяти здоровых и девяти из десяти больных пациентов. Установлено, что в среднем списки биофункций имеют большую степень сходства, чем гены: $\langle J \rangle = 0,63$ для генов и $\langle J \rangle = 0,84$ для биологических функций. Изменения выраженности разных генов в результате могут оказывать влияние на одни и те же биологические функции. Замена пациента в исследуемой выборке вносит в среднем на 20% большую информативность в гены, чем в биофункции.

В случае сравнения двух независимых экспериментальных исследований (образцы взяты у разных людей) были сформированы пары случайных выборок пациентов из доступных данных. В таблице 1 представлены средние рассчитанные индексы Жаккара для списков генов и биологических функций для наборов данных разного размера.

Таблица 1 – Индексы Жаккара в зависимости от объема исследуемой выборки

Размер выборки, шт. здоровые / больные	Средний индекс Жаккара для генов	Средний индекс Жаккара для биофункций
5 / 10	0,19	0,48
5 / 20	0,32	0,55
5 / 40	0,42	0,65
10 / 40	0,42	0,61
10 / 80	0,54	0,66
10 / 100	0,58	0,69
25 / 250	0,76	0,81

При рассмотрении малого количества образцов достоверность результатов анализа достаточно мала, что отражает проблему соотношения результатов независимых исследований. Увеличение размеров выборки повышает идентичность результатов, что свидетельствует о необходимости обмена данными между научными лабораториями.

Необходимо оценить, насколько одно отдельно взятое исследование небольшого количества пациентов будет соотноситься с результатами анализа большого «эталонного» набора данных. Исследовался набор из 550 образцов (50 здоровых и 500 больных). Результаты сравнительного анализа представлены в таблице 2.

Таблица 2 – Индексы Жаккара и косинусные сходства результатов в зависимости от объема исследуемой выборки

Размер малой выборки, шт. здоровые / больные	Средний индекс Жаккара для генов	Средний индекс Жаккара для биофункций	Среднее косинусное сходство для биофункций
5 / 10	0,23	0,19	0,20
5 / 20	0,30	0,25	0,29
10 / 40	0,51	0,57	0,82
10 / 100	0,61	0,68	0,89
25 / 250	0,75	0,84	0,98

Сходство результатов на малых выборках с результатом большой выше, чем при сравнении экспериментов на малых с малыми. Косинусное сходство выше, чем индекс Жаккара для биологических функций, что подтверждает то, что биологические процессы, наиболее сильно подвергающиеся изменениям при болезни, в большей степени схожи, чем те, которые испытывают менее сильное влияние.

ЗАКЛЮЧЕНИЕ

Представленный анализ экспериментальных данных секвенирования РНК позволяет определить дифференциально выраженные гены и биофункции, обогащенные данными генами. Результаты исследований, проводимых на малых выборках, составленных из разных источников экспериментальных данных, имеют низкую степень идентичности. При повышении размера выборки подобие возрастает, что подтверждает необходимость продолжать интеграцию научных исследований и экспериментальных данных геномного секвенирования. Устойчивость списков дифференциально выраженных генов ниже устойчивости биофункций. Для получения статистически значимых результатов достаточно рассмотреть 100 и более раковых образцов.

1. Nazarov, P. V. RNA sequencing and transcriptome arrays analyses show opposing results for alternative splicing in patient derived samples. / P. V. Nazarov [et al.] // BMC Genomics – 2017. – Vol. 18. – Ar. num. 443.
2. Plessis, L. The what, where, how and why of gene ontology – a primer for bioinformaticians / L. Plessis, N. Škunca, C. Dessimoz // Brief Bioinform. – 2011. – Vol. 12. – P. 723–735.
3. Genomic Data Commons [Electronic resource] : National Cancer Institute. – Mode of access: <https://gdc.cancer.gov>. – Date of access: 15.09.2018.
4. Ritchie, M. E. Limma powers differential expression analyses for RNA-sequencing and microarray studies / M. E. Ritchie, B. Phipson // Nucleic Acids Research. – 2015. – Vol. 43. – P. e47.
5. Alexa, A. Enrichment Analysis for Gene Ontology with topGO [Electronic resource] / A. Alexa, J. Rahnenfuhrer // Bioconductor. – Mode of access: <https://bioconductor.org/packages/release/bioc/html/topGO.html> – Date of access: 4.09.2018.
6. Manning, C. D. An Introduction to Information Retrieval Draft / C. D. Manning, P. Raghavan, H. Schütze // Online edition. Cambridge University Press. – 2009. – P. 121.

ПОСТРОЕНИЕ ПОВЕРХНОСТИ ИЗОБРАЖЕНИЯ ПО ПРОИЗВОЛЬНЫМ ЛИНИЯМ

Досин А. Н., Недзьведь А. М., Белоцерковский А. М.
Объединённый институт проблем информатики Национальной академии наук Беларуси
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: nedzveda@tut.by

В данной работе предлагается алгоритм построения поверхности изображения по произвольным линиям на нем, который основан на определении кратчайшего расстояния от каждой точки изображения до ближайшей линии и определения значения на ней. Алгоритм предназначен для реконструкции природных объектов, в частности поверхности горизонтов недр.

ВВЕДЕНИЕ

Задача построения поверхности является одной из самых старых и имеет много решений. Задача является актуальной во многих отраслях, например в сейсмической томографии при построении геологической поверхности, которая основана на суммировании сейсмических трасс, полученных от приемников вдоль линий геологических горизонтов [1,2]. Самый распространенный и эффективный способ построения поверхности для отражения и преломления сейсмического сигнала - это триангуляция. Однако, данный метод имеет ряд недостатков, в частности потерю разрешения и некоторые конфликтные неоднозначности при построении поверхности при помощи треугольников по четырем точкам. В данной работе предлагается алгоритм, позволяющий построить сетку поверхности в необходимом разрешении и разрешить проблемы неоднозначности расположения точек.

I. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ИСХОДНОЙ ИНФОРМАЦИИ

Исходными данными данной задачи являются линии с заданными значениями в каждой точке. Линии пересекаются и могут иметь несколько изломов. Уровень яркости (значения) точек в линиях соответствует либо времени распространения сейсмической волны, либо глубине залегания геологического пласта (рис. 2). Для построения поверхности пересекающейся данные линии, необходимо все данные преобразовать в графическое представление в растровом виде. Таким образом, поверхность строится в виде изображения. Для построения изображения используется диапазон координат в котором лежат графические объекты, который определяет размеры изображения, а для значений пикселей используется вещественный тип, так как нет возможности предсказать значения, которые могут встретиться в числовом ряду. Исходные линии на изображении в каждой точке имеют яркость, соответствующую глубине залегания поверхности.

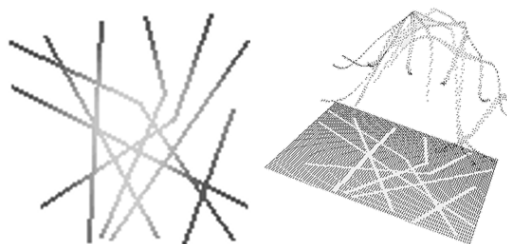


Рис. 1 – Рис. 1 Произвольные линии и их трехмерное представление

II. РАСТРОВОЕ ПОСТРОЕНИЕ ПОВЕРХНОСТИ

Основная идея построения поверхности заключается в определении взвешенной величины от значений пикселей соседних линий по расстоянию от них [3]. Для этого используется карта расстояний [4,5], в которой каждая точка соответствует расстоянию до ближайшей линии и унарная дилатация [5], позволяющая определить близость этой точки к линии. На основе этих данных новое значение глубины определяется по формуле взвешенной суммы:

$$v = v_1 \frac{d_1}{(d_1 + d_2)} + v_2 \frac{d_2}{(d_1 + d_2)},$$

где v – новое значение точки на изображении соответствующее глубине залегания поверхности, v_1 и d_1 – значение ближайшей точки линии и расстояние до неё, v_2 и d_2 – значение ближайшей точки другой соседней линии и расстояние до неё. Таким образом, на основе значений v формируется новое изображение значение которого соответствуют покрывающей поверхности на заданные произвольные линии.

III. ОСОБЕННОСТИ АЛГОРИТМА ПОСТРОЕНИЯ ПОВЕРХНОСТИ

Растровый алгоритм построения поверхности по таким исходным данным включает несколько итераций и специальную структуру представления данных для каждого пикселя. Для оптимизации хранения матриц использовалась специфическая структура пикселя, которая включает следующие переменные:

- значение глубины для первой итерации (v_1);
- значение расстояния для первой итерации (d_1);
- значение глубины для второй итерации (v_2);
- значение расстояния для второй итерации (d_2).

Использование такой структуры в качестве наполнения пиксела позволяет оптимизировать доступ к его элементам матриц, обеспечив компактность расположения элементов в памяти. В результате обработка сводится к построению мультиспектрального изображения. Перед выполнением итераций такое изображение проходит специальную подготовку. Согласно исходным линиям по геометрическим координатам заполняется структура пикселей, где в значения глубины заносится соответствующее значение из линии, а в значение расстояния вносится единица. Таким образом, на изображении формируется исходная информация о линиях. Первая итерация алгоритма представляет собой растровое сканирование изображения слева на право и сверху в низ в ходе которого определяется кратчайшее расстояние от текущей точки до линии в противоположном направлении сканирования, а также происходит рост областей с большими значениями. Во время выполнения шага итерации выполняется анализ окрестности обрабатываемого мультиспектрального изображения. Определяется максимальное значение окрестности, которое переносится в структуру обрабатываемого пиксела в значение глубины для первой итерации. Значение расстояния меняется на максимальное значение расстояния для максимального пиксела из окрестности увеличенное на единицу. Таким образом, наблюдается рост величин в направлении сканирования.

$$v_1 = \max(v_1^{neibor}),$$

$$d_1 = d_1^{\max(neibor)} + 1,$$

где $\max()$ - функция максимума, $neibor$ - соседство или окружение обрабатываемого пиксела. Вторая итерация аналогична первой, однако в данном случае сканирование выполняется справа на лево и снизу в верх для величин v_2 и d_2 . В результате формируются два направления роста величин, которые можно представить в виде четырех матриц: две для определения расстояния по разным направлениям и две для определе-

ния ближайшего значения на линии по двум разным направлениям. Следующая итерация сводит все значения к единому значению в пиксела v . Для каждой отдельной точки это значение пересчитывается по формуле взвешенной суммы, описанной выше. В результате формируется поверхность на основе минимальных расстояний до точек на линиях с максимальными значениями (рис. 2).

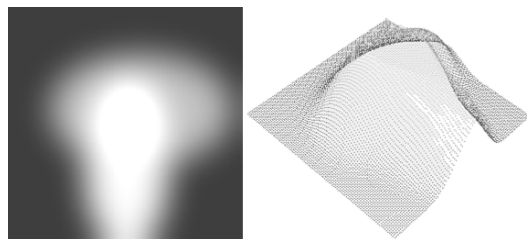


Рис. 2 – Восстановленная поверхность в виде изображения и 3D представлении

IV. ЗАКЛЮЧЕНИЕ

Безусловно такое построение поверхности не является полным соответствием действительности, так как в большинстве случаев данный алгоритм работает в условиях дефицита исходных данных. Однако, правило наименьшего расстояния наиболее эффективно действует при образовании структур в природе. Поэтому алгоритм может быть успешно применен при анализе природных объектов, например, в геологии для томографии. Томография применяется в инженерно-геологических изысканиях для детального изучения механических свойств пород, локализации ослабленных участков, определения степени трещиноватости горных пород, обнаружения подземных пустот естественного и искусственного происхождения.

V. СПИСОК ЛИТЕРАТУРЫ

1. Carcione, J. M. Seismic modeling / J. M. Carcione, G. C. Herman, A. P. E. ten Kroode // GEOPHYSICS Vol 67, № 4, 2002.
2. Бурмин, В. Ю. Методы численного обращения голографов сейсмических волн. М.: Наука, 1993. 100 с.
3. Starovoitov, V. A clustering technique based on the distance transform / V. Starovoitov // Pattern Recognition Letters, DBLP, 1996. – P. 231-239.
4. Недзведь, А. М. Анализ изображений для решения задач медицинской диагностики / А. М. Недзведь, С. В. Абламейко // Минск: ОИПИ НАН Беларуси, 2012. – 239 с.
5. Шапиро, Л. Компьютерное зрение. / Л. Шапиро, Дж. Стокман // изд. — М.: БИНОМ. Лаборатория знаний, 2006. — 752 с.

МЕТОДЫ ОБРАБОТКИ ПОЛЬЗОВАТЕЛЬСКИХ ДАННЫХ ДЛЯ ПРОГНОЗИРОВАНИЯ БИЗНЕС-ПРОЦЕССОВ

Дрозд П. С., Козадаев К. В.

Факультет радиофизики и компьютерных технологий Белорусского государственного университета

Минск, Республика Беларусь

E-mail: drozdps@gmail.com, kozadaeff@mail.ru

В работе представлен новый способ решения задачи сегментации рынка путём кластеризации транзакционных данных клиентов. Данный способ основан на алгоритме локально-чувствительного хеширования, который используется как для снижения размерности исходного пространства признаков, так и для уменьшения количества обрабатываемых образов. Для непосредственной кластеризации значений хеш-функций использовался высокоэффективный алгоритм Bisecting K-mean. Метод показал высокую точность и оказался более устойчивым по сравнению с другими алгоритмами кластеризации. Реализована система, позволяющая в автоматическом режиме сегментировать целевой рынок предприятия.

ВВЕДЕНИЕ

Задача сегментации рынка может быть сформулирована как выделение определённых групп потребителей, для каждой из которых могут потребоваться различные подходы в бизнес-стратегии предприятия. В условиях конкуренции компаниям необходимо определять приоритеты и таргетировать своё предложение на целевые сегменты наиболее перспективных клиентов. В работе [1] решалась задача сегментации рынка с помощью самоорганизующихся карт Кохонена и иерархической кластеризации. Метод показал высокую точность, однако при использовании на практике оказался неустойчивым. Была поставлена цель разработать подход к кластеризации клиентских транзакционных данных, который должен иметь высокую точность и быть более устойчивым.

I. АНАЛИЗИРУЕМЫЙ НАБОР ДАННЫХ

В качестве тестовой выборки для проверки предложенного метода нами был использован набор данных «Ta-Feng», который выложен в свободный доступ компанией ACM RecSys. Он содержит информацию о покупках различных товаров, совершённых более чем 32 тысячами уникальных клиентов. Всего в наборе содержится 817741 запись, каждая из которых описывает совершённую покупателем транзакцию с помощью 9 характеристик (дата проведения платежа, тип товара, сумма транзакции и т.д.). После анализа и предварительной обработки этих данных мы получили $p = 29$ уникальных характеристик для каждой транзакции (этап feature creation).

II. МЕТОД ЛОКАЛЬНО-ЧУВСТВИТЕЛЬНОГО ХЕШИРОВАНИЯ

Алгоритм локально-чувствительного хеширования (LSH - от англ. Locality-Sensitive Hashing) традиционно применяется для решения задачи поиска ближайшего соседа[2], однако нами было принято решение использовать его для

кластеризации. Особенность этого алгоритма заключается в том, что он использует специальный набор «плохих» хеш-функций, которые, в отличие от обычных хеш-функций, должны генерировать коллизии на схожих образах. Таким образом, соседние точки скорее всего попадут в одну хеш-корзину. Пусть $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$ - точка в исходном p -мерном пространстве признаков, которая в рассматриваемом случае описывает одну транзакцию, совершённую одним клиентом (при этом p получилось равным 29 на этапе feature creation). Ограничим значения всех координат этой точки: $C = \sup\{x_1, x_2, \dots, x_p\} + \epsilon, \epsilon \geq 0$ - это условие не накладывает никаких ограничений на специфику решаемой задачи. Преобразуем \mathbf{x} в новый вектор размерности Cp по следующему правилу:

$$\begin{cases} \mathbf{v}(\mathbf{x}) = \Psi_C(x_1)\Psi_C(x_2)\dots\Psi_C(x_p) \\ \Psi_C(x_i) = \underbrace{11\dots11}_{x_i} \underbrace{00\dots00}_{C-x_i}, \forall i \leq p \end{cases}$$

Функция $\Psi_C(x_i)$ переводит значение i -ой компоненты \mathbf{x} в последовательность из x_i единиц, за которыми следуют $C - x_i$ нулей. Например, если $\mathbf{x} = (3, 4)^T$ и $C = 5$, то $\mathbf{v}(\mathbf{x}) = (1110011110)^T$. Хеш-функция в алгоритме LSH вычисляет своё значение для вектора \mathbf{x} путём конкатенации k битов (параметр локально-чувствительного хеширования) из $\mathbf{v}(\mathbf{x})$, порядковые индексы которых содержатся в предварительно сгенерированном множестве Υ , содержащем k случайных целочисленных элементов из $\{1, 2, 3, \dots, Cp\}$. На практике генерируется l множеств $\{\Upsilon_1, \Upsilon_2, \dots, \Upsilon_l\}$ и соответственно l хеш-функций, каждая из которых вычисляет своё значение для каждого вектора исходного набора.

III. АЛГОРИТМ BISECTING K-MEANS

Мы подобрали хеш-функции таким образом, чтобы их значения существовали в метрическом пространстве и между ними можно было вычислять расстояние Хэмминга. После этапа LSH все данные оказались разбиты на доме-

ны, объединяющие схожие транзакции, и каждому домену соответствовало какое-то значение хеш-функции. Для непосредственного кластерного анализа этих значений был применён алгоритм Bisecting K-means - высокоэффективная иерархическая версия K-means. Этот алгоритм является дивизионным, так как предполагает, что все точки изначально принадлежат одному глобальному кластеру[3]. На каждой итерации он делит текущий кластер на два дочерних с помощью обычного K-means с фиксированным параметром $k = 2$. В некоторых случаях [3], Bisecting K-means на порядки производительнее K-means.

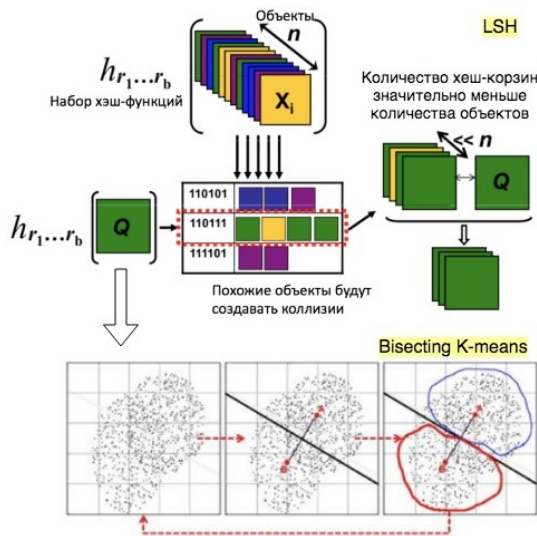


Рис. 1 – Схема предложенного метода

IV. СИСТЕМА ДЛЯ СЕГМЕНТАЦИИ РЫНКА



Рис. 2 – Функциональная схема системы

Предложенный подход был использован при разработке системы для автоматической сегментации рынка. Система основана на стеке технологий Apache® Spark и архитектуре DASE (Data source and preparator, Algorithm, Serving, Evaluation). В качестве клиентского приложения

выступает платформа Salesforce® Sales Cloud, где сохраняются транзакционные данные клиентов после совершения ими каких-либо действий и покупок во внешнем приложении (например, интернет-магазине).

V. РЕЗУЛЬТАТЫ

Результат обработки системой набора данных Ta-Feng представлен ниже.

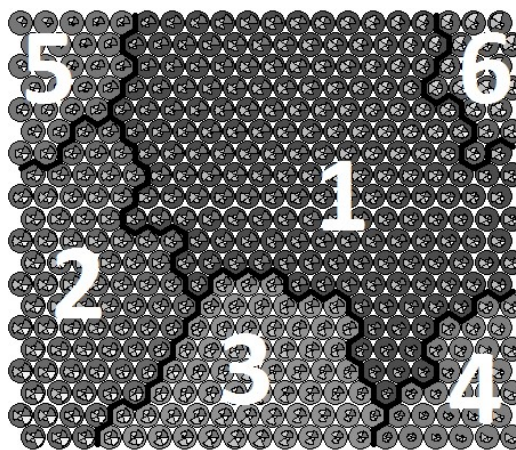


Рис. 3 – Результат сегментации Ta-Feng

Полученная информация может быть использована при планировании маркетинговой стратегии. Например, в кластере 6 набора Ta-Feng находятся VIP-покупатели, совершающие мало очень дорогих покупок. Анализ показал, что выручка от этого сегмента минимальна. Следовательно, бизнесу стоит таргетировать своё предложение на иные группы клиентов.

ЗАКЛЮЧЕНИЕ

Предложенный способ сегментации транзакционных данных показал высокую точность и оказался гораздо более устойчивым, чем классические алгоритмы кластеризации и алгоритмы, основанные на нейронных сетях[1]. Метод имеет высокую производительность, так как используются быстрые приёмы вычисления хеш-функций. Реализована система, автоматизирующая процесс сегментации рынка на основе локально-чувствительного хеширования и алгоритма Bisecting K-means. Система была протестирована на тестовом наборе данных Ta-Feng и в настоящее время проходит апробацию в крупном европейском ритейлере.

1. Drozd P. Kohonen's neural networks for Customer Segmentation / P. Drozd // Open Readings 2018 abstract book / ed. E. Skliutas. – Vilnius, 2018. – P.103.
2. Koga, H. Hierarchical Clustering Algorithm Using Locality-Sensitive Hashing / H. Koga, T.Ishibashi, T. Watanabe // Discovery Science, 7 th International Conference, Padova, 2-5 oct. 2004 / ed. E. Suzuki. – Padova, 2004. – P.114-128
3. Fern, X.Z., Clustering ensembles for high dimensional data clustering / X.Z. Fern, C.E. Brodley // In Proc. International Conference on Machine Learning / ed. T. Fawcett. – Washington DC, 2003. – P.178-185

ВЛИЯНИЕ ПОКАЗАТЕЛЕЙ ИГРЫ ТЕННИСИСТА НА ИСХОД МАТЧА

Канунников И. С., Парамонов А. И.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: i.s.kanunnikov@gmail.com, anton-paramonov@tut.by

В статье рассмотрены спортивные показатели, которые могут влиять на исход матча или турнира. Проведено исследование зависимости результата матча от различных показателей, в частности от характеристик игроков. Выдвинуты предположения о корреляции отдельных показателей игры с результатами, при этом с разной степенью зависимости. Была проведена серия компьютерных экспериментов со статистическими данными, которая подтвердила некоторые гипотезы. Сформулированы цели дальнейших исследований в работе.

ВВЕДЕНИЕ

Для определения зависимости результатов спортивных турниров от различных показателей самой игры необходимо проведение анализа больших объемов статистических данных и построение математических моделей для описания процессов игры.

При рассмотрении игровых результатов в большом теннисе можно констатировать, что исход теннисного матча зависит от многих показателей, как самих игроков, так и всей игры в целом, о чем свидетельствуют профессиональные спортивные источники. Среди таких показателей выделяют: данные о турнире [1] (категория, стадия, характеристики корта [2, 3] и т.п.), данные о конкретном матче (дата, время проведения и т.д.), информация о рейтингах теннисиста [4], статистика личных встреч [5], информация об игроках в динамике (стиль игры, физическое состояние, психологическое состояние и т.д.), мотивация игроков (например, защита рейтинговых очков [6]), эффективность игры теннисиста в данном матче и другие показатели.

На основе обзора источников информации по теннису было выдвинуто предположение, что основное влияние на исход матча имеют показатели эффективности действий игроков.

Эффективность игрока определяется на основе его статистических показателей [7–8], таких как: подача, прием, очки, реализация, выигранные розыгрыши, статистика по типам ударов. Каждый из этих показателей в свою очередь имеет подхарактеристики, которые дают еще более полное представление об игроке.

Для прогнозирования результатов следующего матча (или турнира) в работе предлагается использовать статистические показатели эффективности игры теннисистов в завершенных матчах (в корреляции с исходами этих матчей).

Расчет оценки успешности действий i -го игрока (P_i) выполняется по формуле:

$$P_i = \frac{\sum_{i=1}^m k_i h_i}{m}, \quad (1)$$

где

h_i – характеристика результативности i -го игрока в отдельном игровом компоненте (например количество эйсов за матч);

k_i – весовые коэффициенты, т.к. различные характеристики h_i по-разному влияют на конечный результат;

m – количество характеристик h_i .

В базовой формулировке модели все весовые коэффициенты заданы единичными значениями с целью обнаружения вообще какой-либо зависимости от показателей. В дальнейшем предполагается дополнительное исследование уже уровня значимости тех или иных показателей для определения возможных значений коэффициентов. Это позволит сбалансировать корреляцию показателей с результатов матча.

В результате компьютерных экспериментов по анализу статистики личных встреч нескольких пар игроков были построены графики корреляции показателей действий игроков с исходом конкретного матча.

Были рассмотрены следующие характеристики игрока (h_i) в отдельном матче: попадание 1й подачи, эйсы, выиграно очков на 1й подаче, выиграно очков на 2й подаче, двойных ошибок, винеров, невынужденных ошибок, вынужденных ошибок, агрессия игрока, выиграно очков на приеме, брейкпойнты реализовано.

Анализ данных показал, что в противостоянии игроков все показатели имеют разные влияния на матч, что подтвердило ранее высказанное предположение о необходимости введения коэффициентов полезности.

Выделено три группы воздействия показателей на эффективность игрока: прямое слабое (связь имеется только в нескольких случаях), прямое абсолютное (связь отмечается в абсолютном большинстве исходах) и обратное (отмеча-

ется связь показателя игрока с показателем эффективности противника).

Один из примеров показателей эффективности можно рассмотреть «агрессию игрока» (Aggressive Margin) [9].

Согласно автору методики Хобсону, очки в теннисе могут зарабатывать тремя способами:

1. Собственными ударами навывлет – винарами (winners, W_{own});
2. Вынужденными ошибками соперника (opponent's forced errors, F_{opp});
3. Невынужденными ошибками соперника (opponent's unforced errors, U_{opp}).

Также можно выделить способ «потери очков» (дополнительный способ зарабатывания очков соперником) – собственные невынужденные ошибки (U_{own}), которые снижают собственную эффективность самого теннисиста.

На U_{opp} сам игрок повлиять не может. В то же время, игрок может попытаться увеличить W_{own} , F_{opp} и уменьшить U_{own} .

В общем случае Aggressive Margin теннисиста вычисляется по формуле:

$$A = (W_{own} + F_{opp}) - U_{own}. \quad (2)$$

Пример показателей агрессии игроков в привязке к результатам состязаний отражен на рисунке 1. Как видно, в абсолютном большинстве спаррингов победителем выходил более агрессивный игрок.

Однако преимущество в Aggressive Margin не является прямым залогом победы и его необходимо рассматривать в комбинации с другими показателями.

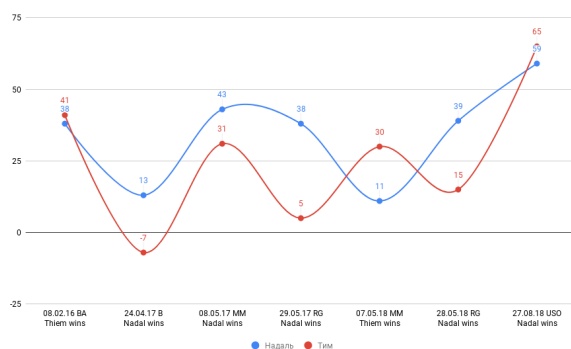


Рис. 1 – Динамика фактора агрессии в личных встречах теннисистов

Таким образом, можно заметить, что отдельные показатели напрямую влияют на эффективность игрока, но при этом косвенно влияют и на показатели эффективности действий соперника.

ЗАКЛЮЧЕНИЕ

На основе полученных данных были даны усредненные оценки количественных показателей для данных статистических метрик, которые должны обеспечить перевес в каждом конкретном компоненте h_i . Результаты данной работы будут полезны теннисистам и тренерам для подготовки и анализа результатов теннисных матчей.

Дальнейшие перспективы использования данной модели заключаются в создании соответствующей специализированной программной системы для спортсменов и тренеров на базе данной модели.

Также данную систему можно будет экстраполировать на родственные теннису виды спорта.

СПИСОК ЛИТЕРАТУРЫ

1. Данные теннисных турниров. Официальный сайт АТР-тура [Электронный ресурс]. - Режим доступа: <https://www.atpworldtour.com/en/tournaments>. - Дата доступа 20.08.2018
2. Данные о скорости кортов. Официальный сайт Международной теннисной федерации [Электронный ресурс]. - Режим доступа: <https://www.itftennis.com/technical/courts/court-testing/court-pace.aspx>. - Дата доступа 20.08.2018
3. Исследование скорости кортов. Электронное хранилище теннисной статистики [Электронный ресурс]. - Режим доступа: <http://www.tennisabstract.com/blog/2011/09/13/the-speed-of-every-surface/> - Дата доступа 20.08.2018
4. Рейтинговая таблица мужского теннисного тура. Официальный сайт АТР-тура [Электронный ресурс]. - Режим доступа: <https://www.atpworldtour.com/en/rankings/singles> - Дата доступа 20.08.2018
5. Статистика личных встреч теннисистов. Новостной портал о теннисе [Электронный ресурс]. - Режим доступа: <http://www.gotennis.ru/players/statistics/ATP/9817> - Дата доступа 20.08.2018
6. Начисление рейтинговых очков за участие в турнирах. Официальный сайт АТР-тура [Электронный ресурс]. - Режим доступа: <https://www.atpworldtour.com/en/rankings/rankings-faq> - Дата доступа 20.08.2018
7. Описание показателей теннисной статистики. Профессиональный трекер теннисной статистики [Электронный ресурс]. - Режим доступа: <http://www.tennis-math.com/tutorial.html1-basic>. - Дата доступа 25.08.2018
8. Статистические показатели матчей теннисиста в сезоне. Электронное хранилище теннисной статистики [Электронный ресурс]. - Режим доступа: <http://www.tennisabstract.com/cgi-bin/player.cgi?p=DominicThiem.amp;q=A2018qq,2017qq,2016qq.amp;q=RafaelNadal>. - Дата доступа 25.08.2018
9. Описание показателя Aggressive margin. Официальный сайт теннисной академии INSPIRATION ACADEMY [Электронный ресурс]. - Режим доступа: <http://hobsontennis.com/blog/the-aggressive-margin/>. - Дата доступа 25.08.2018

РАЗРАБОТКА АЛГОРИТМОВ АНАЛИЗА БИОЧИПОВ ДНК С УЧЕТОМ ВЕСОВЫХ ФАКТОРОВ КАЧЕСТВА СПОТОВ

Климук И.В., Свидрицкий А.С., Яцков Н.Н.

Кафедра системного анализа и компьютерного моделирования, факультет радиофизики и компьютерных технологий, Белорусский государственный университет

Минск, Республика Беларусь

E-mail: ivanklimuk96@gmail.com

В работе предложены три модификации алгоритма k -ближайших соседей для анализа биочипов ДНК с учетом весовых факторов качества спотов. Представлены результаты сравнения алгоритмов k -ближайших соседей без учета и с учетом параметра качества спотов на примерах смоделированных данных.

ВВЕДЕНИЕ

Биочипы или микрочипы ДНК используются для одновременного исследования экспрессии множества генов [1–2]. С помощью технологии биочипов ДНК можно за короткое время обнаружить различные онкологические заболевания, туберкулез и др. [3]. Часто обработка и анализ биочипа затруднены вследствие низкого качества данных и высокого уровня экспериментального шума. Повысить эффективность алгоритмов анализа можно, если учесть параметр качества изображения каждого спота биочипа [4–6]. В работе выбран популярный и универсальный алгоритм классификации k -ближайших соседей.

Цель работы – разработка имитационной модели и алгоритмов k -ближайших соседей для классификации генов с учетом параметра качества спотов на биочипах ДНК.

I. ИМИТАЦИОННАЯ МОДЕЛЬ БИОЧИПА

В основе разработанной имитационной модели лежит алгоритм, представленный в [7]. Алгоритм:

Шаг 1. Задание параметров модели: N – число генов, m – число репликантов, p – доля невыраженных генов в выборке.

Шаг 2. Создание и заполнение матрицы M размером $N * m$: первые $N * p$ строк – значением 0, следующие $\frac{N*(1-p)}{2}$ строк – значением -1, последние $\frac{N*(1-p)}{2}$ строк – значением 1.

Шаг 3. Генерация вектора параметров качества спотов Q размером N с использованием бета-распределения с параметрами $a = 2.5$, $b = 3.5$. Данное распределение наиболее близко имитирует значения реальных экспериментов [5].

Шаг 4. Добавление нормального шума к данным:

$$M_i = M_i + rnorm * (1 - Q_i),$$

где $rnorm$ – реализация стандартной нормальной случайной величины.

Строки матрицы M – объекты или гены для классификации, вектор Q – вектор параметров качества каждого объекта.

II. МОДИФИКАЦИИ МЕТОДА k -БЛИЖАЙШИХ СОСЕДЕЙ

В ходе анализа биочипов ДНК необходимо классифицировать гены на выраженные, невыраженные и подавленные (значение относительной экспрессии 0, 1, -1 соответственно)

Разработаны и программно реализованы 3 модификации метода k -ближайших соседей с учетом весовых факторов качества спотов.

1) kNN с учетом параметра качества спота при расчете расстояния между объектами. Расстояния между объектами обучающей и тестируемой выборки:

$$d_{ij}^w = \frac{d_{ij}}{q_j},$$

где q_j – параметр качества j -го объекта обучающей выборки.

2) kNN с учетом параметра качества спота при голосовании. Параметр качества учитывается при назначении метки класса объекту тестируемой выборки. Класс-победитель выбирается по максимальной сумме их параметров качества:

$$Class(n_i) = arg \max_{c \in C} \sum_{j=1}^l [c_j = c] q_j,$$

3) kNN с учетом параметра качества спота при расчете расстояний между объектами и при голосовании. Алгоритм включает модификации п. 1) и 2)

III. РЕЗУЛЬТАТЫ

Эффективность работы алгоритмов оценивалась как процент ошибки при классификации всех генов, невыраженных и выраженных генов (представляющих наибольший интерес при анализе биочипов). Выполнено сравнение эффективности алгоритмов в зависимости от различных значений параметров обучающей и тестируемой выборки, а именно:

- N_L – размер обучающей выборки, менялся от 100 до 2000;

- p_L – относительная доля невыраженных генов в обучающей выборке, менялось от 0.05 до 0.95;

- $\langle Q \rangle$ и $\langle Q_L \rangle$ – средние значения параметров качества тестируемой и обучающей выборки соответственно, менялись от 0.1 до 0.9.

Зависимость ошибки классификации от параметров модели для разработанных алгоритмов представлены на рисунках 1–4. Модифицированные алгоритмы имеют меньшую ошибку классификации, чем классический алгоритм. На графиках, изображенных на рисунках 1–3, кривая 1 – kNN , 2 – kNN с измененным голосованием, 3 – kNN с пересчетом расстояний и кривая 4 – kNN с измененным голосованием и пересчетом расстояний.

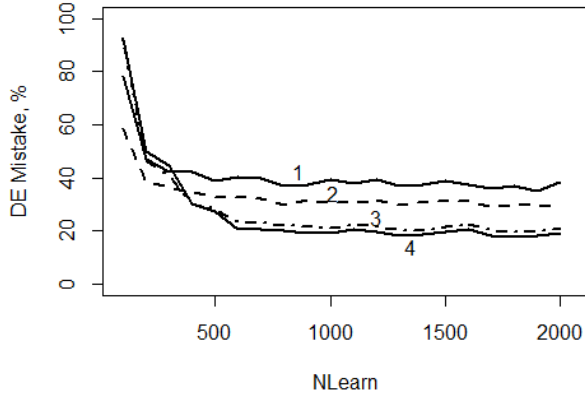


Рис. 1 – Зависимость ошибки классификации выраженных генов от размера обучающей выборки

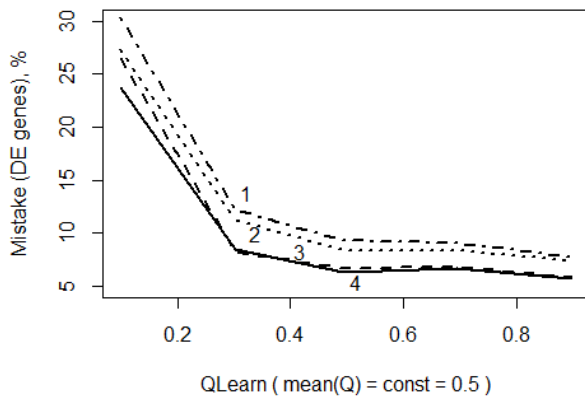


Рис. 2 – Зависимость ошибки классификации выраженных генов от среднего качества тестируемой выборки

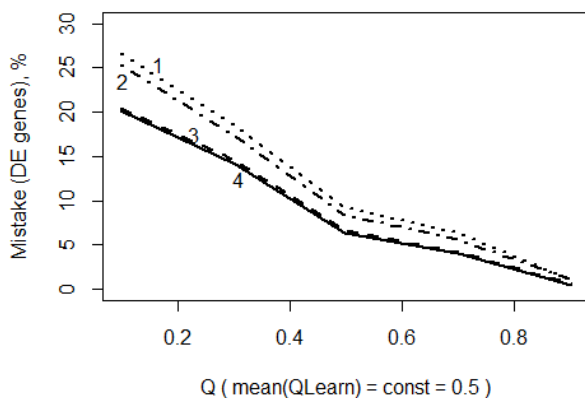


Рис. 3 – Зависимость ошибки классификации выраженных генов от среднего качества обучающей выборки

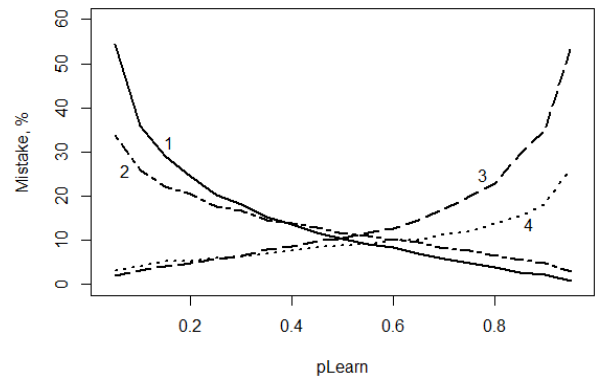


Рис. 4 – Зависимость ошибки классификации от доли невыраженных генов в обучающей выборке:

1 – невыраженных генов для kNN , 2 – невыраженных генов для kNN с измененным голосованием и пересчетом расстояний, 3 – выраженных генов для kNN , 4 – выраженных генов для kNN с измененным голосованием и пересчетом расстояний

IV. ВЫВОДЫ

Разработаны и реализованы метод k -ближайших соседей и три его модификации с учетом параметра качества спотов. Сравнительный анализ эффективности работы алгоритмов на смоделированных данных позволяет сделать выводы:

1) при размере обучающей выборки 500 и более ошибка классификации для всех алгоритмов не меняется;

2) чем хуже качество данных (как обучающей выборки, так и тестируемой), тем лучше с задачей классификации справляются модифицированные алгоритмы;

3) наилучший алгоритм – k -ближайших соседей с учетом параметра качества для расчета расстояний и при голосовании: ошибка классификации на 7% ниже, чем у классического алгоритма.

1. Okuzaki, D. Microarray and whole-exome sequencing analysis of familial Behçet's disease patients/ D. Okuzaki et al – URL://www.ncbi.nlm.nih.gov/pmc/articles/PMC4726226/ (дата обращения: 08.06.2017).
2. Мирзобеков, А. Д. Биочипы в биологии и медицине 21го века //Вестник Российской Академии Наук. 2003. Т. 73. №5. С. 412.
3. Zou, J. Analysis of microarray-identified genes and microRNAs associated with drug resistance in ovarian cancer/ J. Zou, F. Yin, Q. Wang // International Journal of Clinical and Experimental Pathology. 2015.
4. Novikov, E. An algorithm for automatic evaluation of the spot quality in two-color DNA microarray experiments/ E. Novikov, E. Barillot // BMC Bioinformatics. 2005.
5. Yatskou, M. Advanced spot quality analysis in two-colour microarray experiments/ M. Yatskou, E. Novikov, G. Vetter, A. Muller, E. Barillot, L. Vallar, E. Friederich. //BMC Research Notes (2008)
6. Novikov, E. Software package for automatic microarray image analysis (MAIA)/ E. Novikov, E. Barillot //BMC Bioinformatics. 2007. V. 5. P.639.
7. Dembele D. A Flexible Microarray Data Simulation Model // Microarrays. 2013. 2. P. 115-130.

РАЗРАБОТКА ЭВРИСТИЧЕСКОГО АЛГОРИТМА ПОИСКА ОПТИМАЛЬНОГО МАРШРУТА

Козарь Р. В., Навроцкий А. А., Кузнецова О. В., Боброва Т. С.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: pozitr0n.kozarroman@gmail.com, navrotsky@bsuir.by, t.bobrova@bsuir.by, o.kuznetsova@bsuir.by

В данной работе представлена модификация эвристического алгоритма локального поиска для нахождения оптимального маршрута доставки товаров.

ВВЕДЕНИЕ

В настоящее время практически на всех предприятиях существуют очень много проблем, связанных с транспортной логистикой. Основной является проблема контроля своевременной доставки товаров покупателям. Зачастую эта проблема является следствием нестабильной ситуации на дорогах: а именно, заторов и дорожных пробок. Для решения подобного рода проблем в программные продукты автоматизации транспортной логистики, которые используются сотрудниками транспортных отделов, создаются и внедряются алгоритмы, позволяющие строить оптимальный маршрут доставки товаров при различных дорожных ситуациях. Однако транспортных средств с каждым днем становится все больше и вследствие этого существующие алгоритмы становятся все менее оптимальными. Для решения выше описанной проблемы и был создан новый алгоритм поиска оптимального маршрута доставки товаров.

I. ОПИСАНИЕ АЛГОРИТМА

В начале работы алгоритма выбирается начальное решение (найденное любым алгоритмом или выбранное случайно), которое затем итерационно приводится к оптимальному решению. На каждом шаге локального спуска происходит переход от текущего решения к соседнему решению с меньшим значением целевой функции до тех пор, пока не будет достигнут локальный оптимум [1]. Соседнее решение не обязательно должно быть наилучшим в окрестности, но критерий оценки решения не должен меняться во время итеративного процесса. Таким образом, для любого решения s должно быть задано некоторое множество $N(s)$ соседних решений, которое и называется окрестностью s .

II. МАТЕМАТИЧЕСКОЕ ОБОСНОВАНИЕ

Поведение алгоритма зависит от следующих факторов:

1. Структуры окрестностей N . Размер окрестности любого решения должен выбираться на основе компромисса между целью получения хорошего улучшения при каждом переходе к новому решению и целью огра-

ничения времени просмотра одной окрестности. Обычно, для любого решения s , окрестность $N(s)$ порождается с помощью некоторой операции локального изменения s (применительно к текущей ситуации – это абсолютно непредсказуемое изменение дорожной обстановки).

2. Начального решения s_0 . Его можно найти с помощью любого алгоритма (например, конструктивной эвристики), который выдает допустимое решение, или с помощью процедуры случайной генерации значений.
3. Стратегии выбора новых решений. Например, просмотр всех решений из $N(s)$ и выбор наилучшего. Если найденное решение не является оптимальным для локальной области, то осуществляется переход к наилучшему соседу, или к первому лучшему решению, найденному в окрестности. В эвристиках локального поиска используются окрестности, определяемые с помощью последовательностей ограниченной длины операций локального обмена [2]. Пусть, при фиксированном целом $k > 0$, решается этим методом индивидуальная задача проблемы P . Решение u находится в k -обменной окрестности s , если из решения s можно получить решение u , применив не более чем k операций локального обмена. Эвристики, которые основаны на k -обменных окрестностях, часто называют k -оптимальными (k -opt) эвристиками.

III. ДЕТАЛЬНОЕ ОПИСАНИЕ АЛГОРИТМА

Как уже было сказано ранее, новый алгоритм основан на эвристическом алгоритме локального поиска (локального спуска) 3-обменной окрестности вместо 2-обменной окрестности. Данная замена позволяет алгоритму не останавливаться в точке локального оптимума, как это предусмотрено в алгоритме 2-opt, а путешествовать от одного локального оптимума к другому для того, чтобы найти среди всех обходов глобальный оптимум. Для осуществления данного процесса был разработан механизм, позволяющий алгоритму осуществлять выход из локального оптимума (список исключений $List_l(i_k)$). Он

строится по истории поиска, т.е. по нескольким последним точкам $i_k, i_{k-1}, \dots, i_{k-l+1}$ и запрещает исследовать часть окрестности $N(i_k)$ текущего решения i_k . Таким образом на каждом шаге алгоритма очередная точка i_{k+1} является оптимальным решением следующей подзадачи: $m(i_{k+1}) = \min\{m(j) | j \in N(i_k) \setminus List_l(i_k)\}$ при условии, что $List_l(i_k) \subseteq N(i_k)$.

IV. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Проведен тест трех алгоритмов. В качестве среды для анализа был использован Matlab. В качестве функций для тестирования были взяты логарифмическая и линейная функции с различными коэффициентами:

- 1-opt 1-окрестности;
- 2-opt 2-окрестности;
- 3-opt 3-окрестности(новый алгоритм).

Результаты первого теста представлены на рисунке 1.

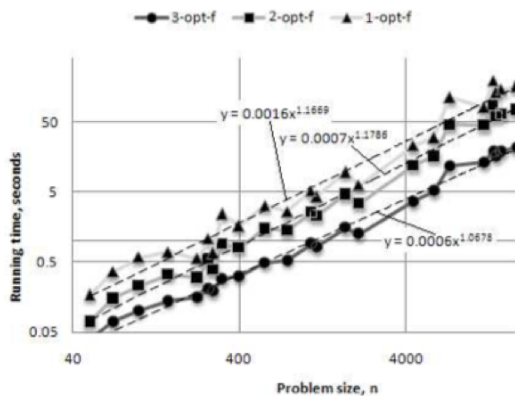


Рис. 1 – Результаты первого теста

Из данных этого графика видно, что производительность 1-окрестного алгоритма весьма низкая, поскольку его временные затраты очень высоки. Данные 2-окрестного алгоритма несколько лучше, но также достаточно затратные. Данные 3-окрестного алгоритма намного лучше, чем 1-окрестный и 2-окрестный алгоритмы, что показывает шкала временных затрат. Так же, для большей наглядности и достоверно-

сти эксперимента было проведено второе тестирование. Результаты второго тестирования алгоритмов поиска оптимального маршрута представлены на рисунке 2.

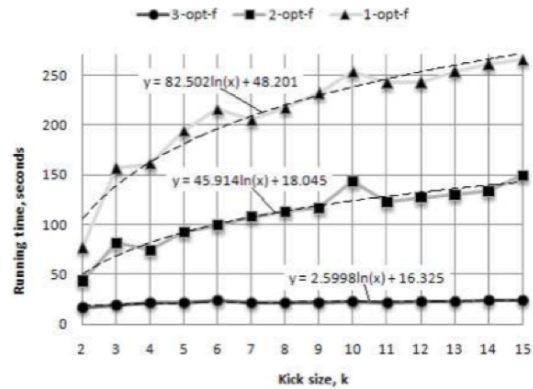


Рис. 2 – Результаты второго теста

Из поставленных экспериментов можно судить, что предложенный в работе алгоритм 3-opt имеет более высокую производительность и более высокое качество полученного результата.

V. ВЫВОДЫ

В работе представлено краткое описание и математическое обоснование нового эвристического алгоритма поиска оптимального маршрута.

1. <https://habrahabr.ru/post/119158/> - Документация по эвристическим и метаэвристическим алгоритмам
2. <https://www.recyclebin.ru/ВМК/П/ii.html> - Использование эвристических алгоритмов в ИИ
3. Ахо, А. В. Введение в структуры данных и алгоритмы / А. В. Ахо, Д. Э. Хопкрофт, Д. Д. Ульман // М., СПб., Киев: «Вильямс», 2016. – 370 с.
4. <https://dic.academic.ru/dic.nsf/ruwiki/703585> - Математический интернет-портал "Академик"
5. <https://habr.com/company/wunderfund/blog/277143/> - Интернет-портал "Хабрахабр". Раздел оптимизации и ускорения эвристических алгоритмов.
6. Гудман, С. С. Введение в разработку и анализ алгоритмов / С. С. Гудман // М.: "Мир" 2010. – 265 с.
7. Карпенко, А. П. Введение в современные алгоритмы поисковой оптимизации / А. П. Карпенко // М.: "Москва" 2007. – 117 с.

ОБ ОСНОВНЫХ АСПЕКТАХ РАЗРАБОТКИ ИНТЕРНЕТ-ПРИЛОЖЕНИЯ «ВИРТУАЛЬНЫЙ ПОМОЩНИК ТУРИСТА»

Кузнецов А. В., Рудикова Л. В., Жавнерко Е. В.

Кафедра современных технологий программирования, Гродненский государственный университет имени
Янки Купалы

Гродно, Республика Беларусь

E-mail: kuzniecov@mail.ru, {rudikowa, eugene.zhavnerko}@gmail.com

В статье описываются общие подходы к разработке системы «Виртуальный помощник туриста». Приводится общая характеристика разрабатываемого приложения, его назначение и средства с помощью которых приложение будет разрабатываться.

ВВЕДЕНИЕ

В современном мире путешествия между странами и внутри стран пользуются большим спросом. Когда границы открыты, а возможности для путешествий ограничены только воображением, очень легко запутаться и предусмотреть не все аспекты своего путешествия. Среди ресурсов Интернета существует большое количество приложений, позволяющих без обращения в туристические агентства, организовать отдых. Однако существенным их недостатком является то, что каждое из них специализируется на одном - покупке билетов, бронирование отелей, автомобилей. Идеальным вариант - собрать весь этот функционал в одном месте, чтобы иметь возможность используя только в одно приложение, полностью спланировать свой отдых. Именно с этой целью и разрабатывается система "Виртуальный помощник туриста". Приложение "Виртуальный помощник туриста" поможет грамотно и, не выходя из дома, спланировать своё путешествие. Оно возьмет на себя такие функции как подбор билетов, отелей, парковки для машины, аренды автомобиля, основываясь на информации, предоставленной в свободном доступе в интернете. Данная система также будет предоставлять информацию о достопримечательностях и интересных местах в выбранных пользователем городах.

I. ЭТАПЫ РАЗРАБОТКИ ПРИЛОЖЕНИЯ

Для обеспечения необходимой функциональности необходимо пройти несколько этапов. В первую очередь, необходимо найти источники данных, на основе которых будет формироваться предоставляемая пользователям информация. В свободном доступе, чаще всего, не предоставляется удобного программного интерфейса для выборки данных, поэтому необходимо выбирать информацию из веб-страниц. Таким образом, вторым шагом является либо поиск универсального средства для разбора веб-страниц, либо написание собственной модуля по конкретные сайты. На третьем шаге собранную информацию необходимо проанализировать и разместить в базу

данных. Результатом работы после третьего шага является схема базы данных, а также модуль, который будет подготавливать информацию для хранения и использования непосредственно приложением. На четвертом шаге разрабатывается серверная часть приложения - подготавливается слой для работы с данными, сервисный слой для форматирования данных и вычисления новых значений на основе существующих данных и слой для формирования ответов. И, наконец, пятым шагом является разработка клиентской части, которая будет заниматься отображением данных для пользователей и обработкой данных, введенных пользователями. При установке приложения происходит создание схемы базы данных и запускается процесс первичного сбора и анализа информации, необходимой для корректной работы приложения. В дальнейшем процесс анализа данных будет проходить в фоновом режиме через определенные промежутки времени.

II. ПОДХОДЫ К РЕАЛИЗАЦИИ АРХИТЕКТУРЫ ПРИЛОЖЕНИЯ

Предлагаемая система представляет из себя веб-приложение. Для его использования необходим только браузер. При реализации приложения было решено разделить функционал на несколько модулей. Это позволит разделить ответственность между несколькими компонентами системы. Первый модуль будет отвечать за разбор веб-страниц, второй модуль за анализ данных полученных при разборе веб-страниц и обработку их, по запросу пользователя, третий - за отображения данных и первичную обработку пользовательской информации. Модуль для разбора веб-страниц - написан на языке Java, использует средства фреймворка Selenide для считывания веб-страниц. При анализе доступной в Интернете информации был выделен набор сайтов с необходимыми данными. Для каждого из таких веб-ресурсов было проанализировано расположение информации на страницах и написан класс для разбора информации. Выбранная информация сохраняется в базу данных для дальнейшего анализа. Данный модуль самостоятелен

относительно основного приложения. Первый запуск данного модуля, как уже упоминалось ранее, происходит при старте приложения. В дальнейшем существует гибкий набор настроек частоты срабатывания выборки данных, в избежании блокировок со стороны ресурсов, из которых производится выборка данных, а также во избежании лишней нагрузки на сервер, на котором будет развернуто данное приложение. Для реализации модуля анализа данных, также был выбран язык Java. Данный модуль выполняет широкий набор задач и является основой системы. Первая задача это разбор информации после ее считывания из веб-страниц. Данные после выборки помещаются в необработанном виде в базу данных. При разборе, данные форматируются под необходимые модели, которые будут использоваться клиентской частью. Когда данные отформатированы они сохраняются в базу данных и готовы для проведения вычислений и отображения на страницах приложения. Второй задачей данного модуля являются вычисления по запросу пользователя. Примером таких вычислений, может быть расчёт стоимости и длительности путешествия между двумя заданными пользователем точками с заданными фильтрами в виде типов использованных транспортных средств и маршрутов. Для таких расчётов применяются вычисления основанные на теории графов. Между заданными точками по существующим в системе данным строится дерево и по необходимому критерию ищется кратчайший путь. Также примером может быть вычисление эмоциональной окраски пользовательских комментариев. С комментарием будет приходить число, показывающее насколько негативен или позитивен данный комментарий. В зависимости от этого модуль отображения информации будет подсвечивать комментарии разными цветами. Для вычисления эмоциональной окраски производится суммирование окраски каждого слова или сочетания слов. Третьей задачей данного модуля является формирование на основе вычислений и существующих данных ответов модулю отображения. Для этого данный модуль предоставляет REST интерфейс (с помощью Spring REST). Данные перемещаются между компонентами в формате JSON. Модуль для отображения или пользовательский интерфейс написан с использованием фреймворка React JS. Данный модуль, используя информацию, предоставленную модулем анализа данных, подготавливает данные к отображению, форматируя их и добавляя разметку. Также данный модуль отображает различный формы и фильтры для пользовательского ввода. На формах присутствует большой набор валидаторов и форматоров, чтобы избежать лишней нагрузки на сервер, в связи с неправильно отформатированными данными. В приложе-

нии также организована ролевая модель. Всего пользователю будет доступно две роли - гость и авторизованный пользователь и одна роль будет для администраторов системы. Для гостевого пользователя будет доступен только просмотр предоставляемой приложением информации и бесплатная регистрация. Для авторизованного пользователя будет доступен следующий набор действий:

- поиск билетов, парковок, аренда автомобилей через сторонние сервисы;
- построение маршрутов поездок;
- публикация маршрутов в приложении, а также возможность написать отзыв, поставить оценку, порекомендовать другим пользователям данный маршрут либо объединить маршруты в тур;
- дневник путешественника - возможность составить список посещенных мест, добавление мест планируемых к посещению, различные заметки, интересные факты о путешествии;
- авторизация через социальные сети и электронную почту.

Администраторы будут иметь следующие возможности:

- все возможности зарегистрированного пользователя;
- добавление и редактирование информации;
- управление пользовательскими профилями (блокировка / разблокировка, сброс пароля);
- доступ к статистической информации о приложении.

III. Выводы

Предлагаемая система будет актуальна в настоящее время. Она будет полезна многим людям, желающим самостоятельно спланировать свой отдых, узнать интересные подробности о достопримечательностях, поделиться с другими людьми своим мнением о путешествии по сред-ствам оценок.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Кузнецов, А.В. О разработке интернет-ресурса, связанного с туристической деятельностью / А.В. Кузнецов // Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях : Материалы XXI Республиканской научной конференции студентов и аспирантов 19-21 марта 2018 г., Гомель. – Гомель : Изд. центр ГГУ им. Ф.Скорины, 2018. – С. 244-245
2. Рудикова, Л.В., Кузнецов, А.В., Жавнерко, Е.В. О разработке интернет-ресурса, связанного с туристической деятельностью / Л.В. Рудикова, А.В. Кузнецов, Е.В. Жавнерко // ИННОВАЦИОННОЕ РАЗВИТИЕ НАУКИ И ОБРАЗОВАНИЯ : сборник статей Международной научно-практической конференции. В 2ч. Ч. 1. – МЦНС «Наука и Просвещение». – 2018. – С. 71-74.

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Кузнецова О. В., Протченко Е. В., Боброва Т. С.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: ovkuznetsova@bsuir.by, protchenko@bsuir.by, t.bobrova@bsuir.by

Использование в медицинской диагностике методов распознавания и создание на их основе автоматизированных систем коренным образом изменяет эффективность диагностики с точки зрения широты и глубины охвата медицинских симптомов, ее оперативности, полноты и достоверности. Одним из наиболее удобных инструментов для решения подобных задач являются искусственные нейронные сети – мощный и одновременно гибкий метод имитации процессов и явлений. Современные искусственные нейронные сети представляют собой программно-аппаратные средства создания специализированных моделей и устройств и позволяют решать широкий круг задач диагностики на основе применения алгоритмов теории распознавания образов. Отличительное свойство нейронных сетей состоит в их способности обучаться на основе экспериментальных данных предметной области.

ВВЕДЕНИЕ

При создании современных информационных систем все активнее осуществляется их интеллектуализация, которая заключается в реализации огромных возможностей разумного поведения живых организмов и человеческого мышления с помощью различных медицинских технологий [1].

Одним из важнейших подходов к интеллектуализации информационных систем является автоматизация процесса распознавания образов [2].

Теория распознавания образов – это раздел кибернетики, развивающий теоретические основы и методы классификации и идентификации предметов, явлений, процессов, сигналов, ситуаций и т.п. объектов, которые характеризуются конечным набором некоторых свойств и признаков. Искусственная нейронная сеть, используемая для распознавания образов, является математической моделью параллельных вычислений, включающей в себя взаимодействующие между собой простые процессорные элементы [3].

I. НЕЙРОННЫЕ СЕТИ В МЕДИЦИНЕ

Применительно к медицинской тематике экспериментальные данные представляются в виде множества исходных признаков или параметров объекта и поставленного на их основе диагноза.

Обучение нейронной сети представляет собой интерактивный процесс, в ходе которого нейронная сеть находит скрытые нелинейные зависимости между исходными параметрами и конечным диагнозом, а также оптимальную комбинацию весовых коэффициентов нейронов, соединяющих соседние слои, при которой погрешность определения класса образа стремится к минимуму [4].

Хорошие результаты показали модели искусственных нейронных сетей для диагностики психических расстройств, сахарного диабета, болезни Паркинсона и Хантингтона. Модели многослойных перцептронов применяются для прогнозирования риска возникновения остеопороза. Логический вывод и обобщенная регрессия использованы для диагностирования гепатита В [5].

II. СУЩНОСТЬ ПРОБЛЕМЫ РАСПОЗНАВАНИЯ

Каждая система распознавания приспособлена для распознавания только данного вида объектов или явлений (так, система, предназначенная для диагностики заболеваний, не может диагностировать отказы аппаратуры, а система, предназначенная для чтения букв русского алфавита, не может читать китайские иероглифы или ноты) [6].

Необходимо рассмотреть содержательную и формальную постановку проблемы распознавания, базирующиеся на следующих положениях. Во-первых, решение задач распознавания требует в общем случае построения специальной системы распознавания. Во-вторых, решение задачи распознавания необходимо (также в общем случае) для того, чтобы система управления, стоящая над системой распознавания, могла принимать правильные решения.

Например, система медицинской диагностики призвана устанавливать диагноз больных для того, чтобы врач мог принимать обоснованные решения о выборе стратегии лечения; система геологической разведки – распознавать наличие и характеристики полезных ископаемых в интересах принятия решений относительно их разработки; система распознавания целей – определять их вид, назначение, характеристики для того, чтобы принимались решения относительно

необходимых мер противодействия этим целям, и т. д.

Исходя из сказанного, системы распознавания должны строиться так, чтобы обеспечивать системе управления возможность наиболее эффективно распоряжаться своими ресурсами, допустимым набором решений, а само построение систем распознавания, как и любых технических систем, не может быть осуществлено без учета соответствующих ограничений.

III. СПЕЦИФИКА ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Широкий круг задач, решаемых с помощью нейросетей, не позволяет пока создать универсальные мощные сети, вынуждая разрабатывать специализированные нейронные сети, функционирующие по различным алгоритмам. Основными преимуществами нейронных сетей для решения сложных задач медицинской диагностики являются:

- отсутствие необходимости задания в явной форме математической модели и проверки справедливости серьезных допущений для использования статистических методов;
- инвариантность метода синтеза от размерности пространства, признаков и размеров нейронных сетей и др.

Однако использование нейронных сетей для задач медицинской диагностики связано также с рядом серьезных трудностей. К ним следует отнести необходимость относительно большого объема выборки для настройки сети, ориентированность математического аппарата на количественные переменные.

Идеальный метод диагностики должен иметь стопроцентную чувствительность и специфичность:

- во-первых, не пропускать ни одного действительно больного человека;
- во-вторых, не переписывать диагноз здоровым людям.

Чтобы предотвратить пропуск случая заболевания, можно обеспечить стопроцентную чувствительность метода. Но в таком случае для системы это оборачивается, как правило, низкой специфичностью метода - у многих людей система будет диагностировать заболевание, которым на самом деле пациент не страдает.

IV. ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Нейронным сетям на этапе обучения необходим «учитель», который заранее выполняет классификацию образов, входящих в обучающую выборку и организует ее ввод в компьютер. В конкурентных сетях реализуется «обучение без учителя». То есть в обучающей выборке

для образов неизвестны правильные (желаемые) выходные реакции. Для реализации этого подхода необходимо решить две основные проблемы:

- разработать методы разбиения образов на классы без учителя – этап обучения;
- выработать правила отнесения текущего входного образа к некоторому классу – этап распознавания.

Обучение нейронной сети в большинстве случаев представляет собой автоматизированный процесс, в котором только после его окончания требуется участие специалиста для оценки результатов. Естественно, часто может потребоваться корректировка, создание дополнительных сетей с другими параметрами и т.д., однако всегда есть возможность оценить работу системы на любом этапе обучения, протестировав контрольную выборку.

Как отмечалось выше, обучение НС выполняется фактически за два этапа:

1. Обучение на обучающем множестве пока не выполнено одно из условий останова:
 - ошибка на обучающем множестве становится меньше заданной величины;
 - ошибка перестает изменяться в течение определенного числа итераций;
 - достигнута верхняя граница числа итераций обучения.
2. Проверка правильности функционирования сети на тестовом множестве.

V. ВЫВОДЫ

Нейросетевые технологии призваны решать трудноформализуемые задачи, к которым, в частности, сводятся многие проблемы медицины.

Отличие методологии проектирования нейросетевых систем от традиционных состоит именно в том, что система никогда не создается сразу готовой, и никогда не является полностью законченной, продолжая накапливать опыт в процессе эксплуатации.

VI. СПИСОК ЛИТЕРАТУРЫ

1. Горелик, А. Л. Методы распознавания / А. Л. Горелик, В. А. Скрипкин // – 2000. – С. 20.
2. Кутыркин, А. В. Распознавание оптических образов (символов) с помощью хемминговой меры близости / А. В. Кутыркин, А. В. Семин // – 2005. – С. 11.
3. Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, А. Джин Понс // – 2004. – С. 928.
4. Яхьяева, Г. Э. Нечеткие множества и нейронные сети / Г. Э. Яхьяева // – 2006. – 316с.
5. Basit, A. Artificial Neural Network: A Tool for Diagnosing diseases / A. Basit, K. Raffat // – 2013. – P. 27–30.
6. Нейронные сети для решения задач диагностики [Электронный ресурс] / –Режим доступа: <https://studopedia.ru/neyronnie-seti-dlya-zadach-diagnostiki.html>. – Дата доступа: 24.09.2018.

ОПТИМИЗАЦИЯ ВЫЧИСЛЕНИЯ ДВУМЕРНЫХ РЕКУРРЕНТНЫХ СВЕРТОК ПРИ ПОМОЩИ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Липницкий В. А., Сергей А. И.

Кафедра высшей математики, Военная академия Республики Беларусь

Факультет экономики и управления, Гродненский государственный университет имени Янки Купалы

Минск, Гродно, Республика Беларусь

E-mail: valipnitski@yandex.by, sergej.a.i@mail.ru

В статье приводится эффективный алгоритм решения одной из подзадач третьей проблемы Кэмерона. Рассматривается применение быстрого преобразования Фурье для вычисления двумерных сверток с рекуррентными зависимостями. Предлагаемый алгоритм снижает асимптотическую сложность расчетов с $O(n^4)$ (для базового алгоритма) до $O(n^{2.5} \log n)$ арифметических операций.

ВВЕДЕНИЕ

Рассматриваемая задача возникает при подсчете количества квадратных бинарных матриц n -го порядка, содержащих в точности n единиц, с точностью до перестановки строк/столбцов (третья проблема Кэмерона [1]).

Наибольшего прогресса в решении проблемы удалось достичь при помощи метода динамического программирования [2]. При таком подходе естественным образом возникает необходимость вычисления сложных сверток, которые в свою очередь являются узким местом алгоритма. Поэтому оптимизация расчетов приведенных ниже сверток напрямую влияет на итоговую сложность алгоритма решения проблемы Кэмерона.

I. ПОСТАНОВКА ЗАДАЧИ

Необходимо эффективно вычислить значение величины

$$r_{i,j} = \frac{1}{i} \sum_{k=0}^{i-1} \sum_{l=0}^j r_{k,l} t_{i-k,j-l} \quad (1)$$

для $1 \leq i \leq n, 0 \leq j \leq n$ по заданным $r_{0,j}, 0 \leq j \leq n$ и $t_{i,j}, 0 \leq i, j \leq n$.

Соотношение (1) напоминает формулу для вычисления дискретной двумерной свертки [3]. Основное отличие и сложность формулы (1) заключается в рекуррентной зависимости между значениями $r_{i,j}$.

Формула (1) отчетливо показывает двумерную структуру задачи, поэтому ее естественно рассматривать в матричном виде. Пусть

$$R = (r_{i,j})_{\substack{i=0,\overline{n} \\ j=0,\overline{n}}}, T = (t_{i,j})_{\substack{i=0,\overline{n} \\ j=0,\overline{n}}}.$$

Тогда по заданной матрице T и 0-й строке матрицы R нужно вычислить остальные элементы матрицы R по формуле (1).

Оценим асимптотическую сложность расчета значений $r_{i,j}$ пользуясь формулой (1) напрямую. Матрицу R можно вычислять построчно, поскольку элементы i -й строки зависят только от элементов из предыдущих строк.

Сложность вычисления отдельно взятого элемента $r_{i,j}$ составляет $O(ij)$. Следовательно, вычисление всех недостающих элементов матрицы R «наивным» способом имеет сложность $\sum_{i=1}^n \sum_{j=0}^n O(ij) = O(n^4)$.

II. ОПТИМИЗАЦИЯ 1. ОДНОМЕРНЫЕ СВЕРТКИ

Будем оптимизировать вычисление $r_{i,j}$, сводя его к расчету одномерных сверток.

Пусть u и v – вектор-строки длины $n+1$: $u = (u_0 u_1 \dots u_n), v = (v_0 v_1 \dots v_n)$. Тогда их одномерной сверткой будем называть вектор-строку $conv1(u, v)$ длины $n+1$, j -й элемент которой вычисляется по формуле

$$[conv1(u, v)]_j = \sum_{l=0}^j u_l v_{j-l}, j = \overline{0, n}. \quad (2)$$

Пусть $R_i = (r_{i0} r_{i1} \dots r_{in})$ – i -я строка матрицы R и, аналогично, $T_i = (t_{i0} t_{i1} \dots t_{in})$ – i -я строка матрицы T . Тогда Формулу (1) можно записать в виде:

$$R_i = \frac{1}{i} \sum_{k=0}^{i-1} conv1(R_k, T_{i-k}), i = \overline{1, n}. \quad (3)$$

Вычисление свертки $conv1(R_k, T_{i-k})$ по формуле (2) потребует $O(n^2)$ операций, однако быстрое преобразование Фурье позволяет вычислять такие свертки со сложностью $O(n \log n)$ [3]. Таким образом, вычисление R_i по формуле (3) имеет сложность $O(in \log n)$, а общая сложность решения исходной задачи составляет $O(n^3 \log n)$.

III. ОПТИМИЗАЦИЯ 2. ДВУМЕРНЫЕ СВЕРТКИ

Выразим вычисление $r_{i,j}$ через расчет двумерных сверток. Пусть A и B – квадратные матрицы порядка $n+1$. Тогда их двумерной сверткой будем называть матрицу $conv2(A, B)$ порядка $n+1$, элемент $a_{i,j}$ которой вычисляется по формуле

$$[conv2(A, B)]_{i,j} = \sum_{k=0}^i \sum_{l=0}^j a_{i,j-k-l} b_{k,l}. \quad (4)$$

Формулу (4) можно также выразить через применение одномерных сверток к строкам:

$$[\text{conv}2(A, B)]_i = \sum_{k=0}^i \text{conv}1(A_k, B_{i-k}). \quad (5)$$

Рассмотрим способ, позволяющий вычислить очередную строку матрицы R целиком при помощи двумерных сверток. Обозначим через $R^{(i)}$ матрицу, полученную из R занулением строк с $(i+1)$ -й по n -ю, т. е. матрицу $(R_0 \ R_1 \ \dots \ R_{i-1} \ R_i \ 0 \ \dots \ 0)^T$.

Пользуясь формулами (5) и (3), получаем,

$$\begin{aligned} [\text{conv}2(R^{(i-1)}, T)]_i &= \sum_{k=0}^i \text{conv}1([R^{(i-1)}]_k, T_{i-k}) = \\ &= \underbrace{\text{conv}1([R^{(i-1)}]_i, T_0)}_{=(0 \ \dots \ 0)} + \sum_{k=0}^{i-1} \text{conv}1(R_k, T_{i-k}) = iR_i. \end{aligned}$$

Поэтому формула для вычисления i -й строки матрицы R на основе предыдущих имеет вид:

$$R_i = \frac{1}{i} [\text{conv}2(R^{(i-1)}, T)]_i. \quad (6)$$

Вычисление свертки $\text{conv}2(R^{(i-1)}, T)$ по формуле (4) потребует $O(n^4)$ операций, однако и в двумерном случае быстрое преобразование Фурье позволяет ускорить расчет сверток до $O(n^2 \log n)$ [3]. Таким образом, вычисление R_i по формуле (6) имеет сложность $O(n^2 \log n)$, а общая сложность решения исходной задачи таким способом составляет $O(n^3 \log n)$.

IV. ОПТИМИЗАЦИЯ 3. ЧЕРЕДОВАНИЕ ОДНОМЕРНЫХ И ДВУМЕРНЫХ СВЕРТОК

В результате оптимизаций 1 и 2 получились два независимых алгоритма решения исходной задачи, оба со сложностью $O(n^3 \log n)$. В этом разделе эти два метода объединяются для получения более эффективного алгоритма.

Обозначим $\text{conv}1(R_k, T_{i-k})$, промежуточную величину из оптимизации 1, через $C_{i,k}$, $i = \overline{1, n}$, $k = \overline{0, i-1}$. В этом разделе основное внимание уделяется именно значениям $C_{i,k}$ и их вычислению.

Также введем обозначение $P(i)$ для величины $\text{conv}2(R^{(i-1)}, T)$, возникающей в оптимизации 2, $i = \overline{0, n}$. В соответствии с формулой (5), имеем

$$[P(i)]_i = \sum_{k=0}^{i-1} C_{i,k}. \quad (7)$$

Заметим, что следующие за i -й строки матрицы $P(i)$ также несут в себе полезную информацию, например:

$$[P(i)]_{i+1} = \sum_{k=0}^{i-1} C_{i+1,k}, [P(i)]_{i+2} = \sum_{k=0}^{i-1} C_{i+2,k}.$$

В общем случае

$$[P(i)]_{i+s} = \sum_{k=0}^{i-1} C_{i+s,k}, s = \overline{0, n-i}.$$

Отсюда следует, что для $0 \leq s \leq n-i$

$$[P(i+s)]_{i+s} = [P(i)]_{i+s} + \sum_{k=i}^{i+s-1} C_{i+s,k}. \quad (8)$$

Другими словами, имея вычисленное значение матрицы $P(i)$, можно получить значение

$(i+s)$ -й строки матрицы $P(i+s)$, посчитав дополнительно s недостающих значений $C_{i+s,k}$, для $k = \overline{i, i+s-1}$.

Из формулы (6) следует, что $R_i = \frac{1}{i} [P(i)]_i$, $i = \overline{1, n}$, поэтому для решения исходной задачи достаточно вычислить $[P(i)]_i$ для $i = \overline{1, n}$.

Сложность вычисления значения $C_{i,k}$ при помощи одномерной свертки составляет $O(n \log n)$, тогда как сложность вычисления матрицы $P(i)$ при помощи двумерной свертки равна $O(n^2 \log n)$.

Будем вычислять строки матрицы R блоками. Выберем некоторое целое число b - размер блока, $1 \leq b \leq n+1$. Строки $1, 2, \dots, b-1$ матрицы R вычислим, пользуясь формулой (7). Остальные строки будем вычислять следующим образом. Для строк с номерами вида rb будем находить значение матрицы $P(rb)$ и использовать ее при вычислении строк с номерами $rb+s$, $s = \overline{1, b-1}$, используя формулу (8).

На рисунке 1 показан принцип вычисления строк матрицы R по блокам на примере $n = 11, b = 4$. Цвет ячейки (i, k) показывает вычислялось ли значение $C_{i,k}$ отдельно (белый) или было учтено в первом слагаемом формулы (8) (черный).

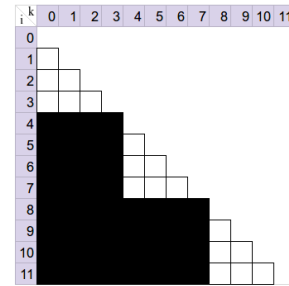


Рис. 1 - Вычисление $C_{i,k}$ блоками для $n = 11, b = 4$

Заметим, что при $b = n+1$ этот алгоритм вырождается в оптимизацию 1, а при $b = 1$ - в оптимизацию 2.

Выберем оптимальный размер блока b . Сложность вычисления матрицы $P(rb)$ составляет $O(\frac{n}{b} b^2 n \log n)$. Сложность подсчета значений $C_{i,k}$ равна $O(\frac{n}{b} n^2 \log n)$. Общая сложность алгоритма: $O(bn^2 \log n + \frac{n^3}{b} \log n)$. Легко видеть, что оптимально выбирать b порядка \sqrt{n} , в этом случае сложность алгоритма получится $O(n^{2.5} \log n)$.

1. Cameron P. J. Problems on permutation groups // P. J. Cameron - [Электронный ресурс] - Режим доступа: <http://www.maths.qmul.ac.uk/~pjc/pgprob.html>. - Дата доступа: 07.02.2017.
2. Липницкий, В. А. Алгоритм развертки в подсчете количества орбит кэмероновских матриц / В. А. Липницкий, А. И. Сергей, Н. В. Спичекова // Веснік Магілёўскага дзяржаўнага ўніверсітэта імя А.А. Куляшова, 2017, №2(50). - С. 30 - 43.
3. Быстрое преобразование Фурье и алгоритмы вычисления сверток / Г. Нуссбаумер; Изд-во: М.: Радио и связь, 1985 г. - 245 с.

ТЕСТИРОВАНИЕ СКАЧКОВ ИЗМЕНЕНИЯ ХАРАКТЕРИСТИК СИГНАЛОВ С ПОМОЩЬЮ ВЕЙВЛЕТОВ

Лобач С. В.

Кафедра математического моделирования и анализа данных,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: lobachS@bsu.by

Применяется вейвлет-анализ для построения тестирующей статистики обнаружения изменений характеристик последовательности независимых случайных величин. Методом статистических испытаний оценивается мощность теста.

ВВЕДЕНИЕ

Анализ изменений (в том числе и скачкообразных) характеристик сигнала привлекает внимание многих исследователей. Впервые задача обнаружения "разладки" была сформулирована в 50-е годы XX века в работе [1]. Первоначально эта задача получила применение в промышленном производстве, медицинских исследованиях, геофизике, задачах технической диагностики, обработке сигналов. Таким образом, класс задач обнаружения разладки является очень широким. Эти задачи отличаются одна от другой предположениями о модели наблюдаемого процесса и подходами к ее решению. Существует два основных подхода к решению задачи разладки: методы апостериорного обнаружения и последовательные методы обнаружения. В первом случае предполагается, что в имеющейся последовательности наблюдений в некоторый момент произошло изменение характеристик и на основе полученных наблюдений необходимо оценить момент изменений. При этом свойства полученных оценок изучаются в асимптотической постановке при объеме наблюдений, стремящимся к бесконечности [2]. При последовательном обнаружении на каждом шаге при поступлении нового наблюдения гипотеза о наступлении разладки либо принимается и наблюдения прекращаются, либо отклоняется и наблюдения продолжают дальше [3].

Одним из основных методов решения задачи обнаружения разладки является метод кумулятивных сумм (CUSUM) [4], причем для изучения CUSUM-статистики используют метод Фурье-анализа [5]. В данной работе применяется вейвлет-анализ для тестирования наличия моментов изменения свойств случайной последовательности случайных величин, которая может служить простейшей математической моделью исследуемого сигнала.

I. ПОСТАНОВКА ЗАДАЧИ

Пусть наблюдения x_1, x_2, \dots, x_n являются независимыми случайными величинами, распре-

деленными по нормальному закону распределения вероятностей:

$$x_i \sim N\left(f\left(\frac{i}{n}\right), \sigma^2\right), \quad (1)$$

где $f(\cdot)$ – известная функция, определенная на отрезке $[0, 1]$. В задаче обнаружения момента изменения среднего значения последовательности независимых случайных величин (1) предполагается, что

$$f(t) = \begin{cases} \mu, & 0 \leq t < \tau, \\ \mu + \Delta, & \tau < t \leq 1, \end{cases} \quad (2)$$

где τ – неизвестный момент времени. Задача состоит, во-первых, в тестировании факта существования момента изменения среднего значения. Если принимается гипотеза о наличии этого момента, то дальнейшее исследование заключается в оценивании параметров $\tau, \mu, \Delta, \sigma^2$. Следует отметить, что ключевой проблемой здесь является оценивание параметра τ . Задача (2) может быть обобщена, если функцию $f(t)$ представить в виде:

$$f(t) = \begin{cases} g(t), & 0 \leq t < \tau, \\ h(t), & \tau \leq t < 1, \end{cases}$$

где $h(\tau_+) - g(\tau_-)$ – скачок среднего.

II. ВЕЙВЛЕТЫ И V-ТЕСТ

Вейвлет-анализ основан на разложении исследуемой функции в ряд по базисным функциям и изучении коэффициентов разложения. Он является удобным аппаратом для исследования разрывов функций, максимумов, минимумов и других особенностей. Удобно использовать объем наблюдений $n = 2^J$, где J – некоторое натуральное число. Вейвлеты представляют собой функции [6]:

$$\{\psi_{jk}(t), k=0, \dots, 2^j - 1, j=0, 1, \dots, J - 1\}, \quad (3)$$

где $\psi_{jk}(t) = 2^{-j/2}\psi(2^{-j}t - k)$.

Функция $\psi(t)$ называется материнским вейвлетом. В случае вейвлета Хаара [6]:

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 0,5, \\ -1, & 0,5 \leq t < 1, \end{cases}$$

Вейвлет-преобразование, примененное к последовательности $\{x_i\}, i = 1, 2, \dots, n$, приводит к рассмотрению вейвлет-коэффициентов

$$d_{jk} = \sum_{i=0}^n x_i \psi_{jk}(t_i). \quad (4)$$

Чтобы изучить вероятность ошибок второго рода используемого статистического критерия, введем последовательность альтернативных гипотез H_{1n} и основную H_0 :

H_0 : x_1, \dots, x_{τ_n} – последовательность независимых случайных величин, распределенных по нормальному закону $N(\mu, 1)$.

H_{1n} : x_{τ_n+1}, \dots, x_n – последовательность независимых случайных величин, распределенных по нормальному закону $N(\mu + \Delta_n, 1)$, $\Delta_n = O(\sqrt{\log_2 n/n})$.

Поскольку наблюдения x_i независимы и распределены по нормальному закону, то коэффициенты d_{jk} , определенные по формуле (4), также будут независимы и нормально распределенными в силу линейности формулы (4) и ортогональности базиса $\{\psi_{jk}(t)\}$ (3). При условии выполнения нулевой гипотезы H_0 коэффициенты $\{d_{jk}\}$ (4) будут независимы и распределены по закону $N(0, 1)$. При условии выполнения альтернативной гипотезы H_{1n} вейвлет-коэффициенты будут также распределены по нормальному закону, но с ненулевым средним.

Используя эти вероятностные свойства вейвлет-коэффициентов, построим статистику критерия:

$$V = \sum_{j=1}^{J-1} V_j,$$

где $V_j = \sqrt{2^{j-1}} \cdot \sum_{k=0}^{2^j-1} d_{jk}$.

Очевидно, при выполнении гипотезы H_0 , статистика критерия $V \sim N(0; \tau \cdot n)$. Таким образом, тест проверки гипотезы H_0 имеет вид:

$$d = \begin{cases} H_0, & \frac{|T|}{\sqrt{\tau \cdot n}} < \Delta, \\ H_1, & \frac{|T|}{\sqrt{\tau \cdot n}} \geq \Delta, \end{cases}$$

где $\Delta = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right)$, где ε – уровень значимости V -критерия; $\Phi(t)$ – функция распределения вероятностей стандартного нормального закона. Вероятность ошибки второго рода вычисляется по формуле

$$w_n = P \left\{ \frac{|T|}{\sqrt{\tau \cdot n}} = \Delta | H_{1n} \right\}.$$

III. КОМПЬЮТЕРНЫЕ ЭКСПЕРИМЕНТЫ

Для оценивания мощности теста проводилась серия компьютерных экспериментов. Объемы выборок составляли 16, 32, 64, 128, 256. Предполагается, что из параметров модели $(\tau, \mu, \Delta, \sigma^2)$ неизвестным является только τ . Для каждой выборки размер скачка составлял $\Delta_n = \sqrt{\log_2 n/n}$. Момент наступления скачка определялся по формуле $\tau_n = [n \cdot \lambda]$, где $[a]$ – целая часть числа a . В качестве значений λ брались 0, 1, 0, 2, 0, 3, 0, 5. Результаты моделирования приведены в таблице (уровень значимости $\varepsilon = 0,05$; $\lambda = 0, 1$):

n	16	32	64	128	256
w	0,945	0,94	0,9	0,94	0,93

Результаты моделирования показывают эффективность предложенной статистики для обнаружения изменений среднего значения случайной последовательности, поскольку мощность критерия близка к 1.

1. Page, E. S. Continuous inspection schemes / E. S. Page // *Biometrika*. – 1954. Vol. 41, №. 1. P. 100-115.
2. Page, E. S. Estimating the point of change in a parameter occurring at an unknown point / E. S. Page // *Biometrika*. – 1955. Vol. 42, №. 4. P. 523-527.
3. Никифоров, И. В. Последовательное обнаружение изменения свойств временных рядов / И. В. Никифоров // М.: Наука. – 1983, 199 с.
4. Gilchrist, C. Note on distribution-free CUSUM technique / C. Mc Gilchrist, K. Woodyer // *Technometrics*. – 1975. Vol. 17, №. 1. P. 321-325.
5. Lombard, F. Detecting Change Points by Fourier Analysis / F. Lombard // *Technometrics*. – 1985. Vol. 30, №. 3. P. 305-310.
6. Смоленцев, Н. К. Введение в теорию вейвлетов / Н. К. Смоленцев // М.-Ижевск: НИЦ "Регулярная и хаотическая динамика". – 2010, 292 с.

ВЕЙВЛЕТ-АНАЛИЗ В ИССЛЕДОВАНИИ НЕОДНОРОДНОСТЕЙ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Лобач В. И.

Кафедра математического моделирования и анализа данных,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: lobach@bsu.by

Предлагается и исследуется статистический критерий обнаружения неоднородностей случайных последовательностей на основе вейвлета Хаара.

ВВЕДЕНИЕ

Информация и изучение ее свойств представляют огромную важность в мире современных компьютерных технологий. Все более широкое распространение при обработке информационных массивов данных приобретают методы вейвлет-анализа [1–3]. Вейвлеты в последнее десятилетие нашли также широкие применения в обработке сигналов и изображений. Теория вейвлетов является альтернативной анализу Фурье и дает более гибкую технику обработки сигналов. Одно из основных преимуществ вейвлет-анализа заключается в том, что он позволяет заметить хорошо локализованные свойства сигнала. Анализ Фурье этого не дает, так как в коэффициентах Фурье отражается поведение сигнала за все время его существования. Целью данной работы является обнаружение неоднородностей случайных последовательностей при помощи данных методов.

I. ВЕЙВЛЕТ ХААРА

Определение 1. Функция $\psi \in L^2(R)$ называется ортогональным вейвлетом, если семейство $\{\psi_{j,k}\}$, определенное формулой $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, $j, k \in Z$, является ортонормированным базисом в $L^2(R)$; это означает, что

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \cdot \delta_{k,m}, \quad j, k, l, m \in Z,$$

и любая функция $f \in L^2(R)$ может быть представлена как

$$f(t) = \sum_{j,k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(t), \quad (1)$$

где ряд (1) сходится в $L^2(R)$, а именно

$$\lim_{M_1, N_1, M_2, N_2 \rightarrow \infty} \left\| f - \sum_{j=-M_2}^{N_2} \sum_{k=-M_1}^{N_1} c_{j,k} \psi_{j,k} \right\|_2 = 0.$$

Ряды, представляющие функции f в (1), называются *вейвлет-рядами*. Вейвлет-коэффициенты $c_{j,k}$ определяются формулой

$$c_{j,k} = \langle f, \psi_{j,k} \rangle.$$

Определим интегральное преобразование W_ψ в $L^2(R)$ как

$$(W_\psi f)(b, a) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad f \in L^2(R),$$

то вейвлет-коэффициенты в (2) принимают вид

$$c_{j,k} = (W_\psi f)\left(\frac{k}{2^j}, \frac{1}{2^j}\right).$$

Рассматривается случайная последовательность:

$$\Xi = \{\xi_t | \xi_t \in \{0, 1\}, t = \overline{0, N-1}\} \quad (2)$$

Положим, не теряя общности, $N = 2^n$, где $n \in N$.

Нулевая гипотеза о распределении вероятностей данной последовательности формулируется в виде:

H_0 : $\{\xi_t\}$ – последовательность случайных, независимых и одинаково распределенных величин. Альтернативная гипотеза:

H_1 : $\xi_t = \eta_t \oplus a_t$, $t = \overline{0, 2^n - 1}$, где $A = \{a_t | a_t \in \{0, 1\}, t = \overline{0, 2^n - 1}\}$ – периодическая с периодом T последовательность детерминированных величин;

$G = \{\eta_t | \eta_t \in \{0, 1\}, t = \overline{0, 2^n - 1}\}$ – последовательность независимых случайных величин, распределенных по закону Бернулли $Bi(1, p)$, причем $p \ll 1/2$.

Определение 2. Дискретное вейвлет-преобразование случайной последовательности (2) задается путем вычисления коэффициентов:

$$d_{j,k}^{(\psi)} = \sum_{t=0}^{2^n-1} \xi_t \psi_{j,k}(t).$$

В качестве вейвлет-функции выберем вейвлет Хаара :

$$\psi^H(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t < 1, \\ 0, & \text{иначе.} \end{cases}$$

Для последовательности (2) и дискретного вейвлет-преобразования Хаара детализирующие

(уточняющие) коэффициенты преобразования имеют дискретное распределение вероятностей. Найдем это распределение вероятностей.

Ортонормированная система вейвлетов Хаара выглядит следующим образом:

$$\psi_{j,k}^H(t) = \begin{cases} 2^{-\frac{j}{2}}, & 2^j k \leq t < (2^j k + 2^{j-1}), \\ -2^{-\frac{j}{2}}, & (2^j k + 2^{j-1}) \leq t < (2^j k + 2^j), \\ 0, & \text{иначе.} \end{cases}$$

Можно показать, что множеством значений $d_{j,k}^{(\psi)}$ является $\pm 2^{\frac{j}{2}-1}; \pm (2^{\frac{j}{2}-1} - 2^{-\frac{j}{2}}); \pm (2^{\frac{j}{2}-1} - 2 \cdot 2^{-\frac{j}{2}}); \dots; \pm (2^{\frac{j}{2}-1} - (2^{j-1} - 1) \cdot 2^{-\frac{j}{2}}); 0$.

Т. е. окончательно можно записать:

$$d_{j,k}^{(\psi)} = \pm (2^{\frac{j}{2}-1} - i \cdot 2^{-\frac{j}{2}}), \quad i = \overline{0, 2^{j-1}}, \quad j = \overline{1, n}.$$

II. РЕЗУЛЬТАТЫ КОМПЬЮТЕРНЫХ ЭКСПЕРИМЕНТОВ

Приведем теорему, которая отражает свойства вейвлет-коэффициентов и является основой для построения статистического χ^2 -критерия проверки гипотезы об однородности исследуемой случайной последовательности.

Теорема [4]. Пусть элементы последовательности (2) $\Xi = \{\xi_t | \xi_t \in \{0, 1\}, t = \overline{0, 2^n - 1}\}$ независимы и одинаково распределены $P\{\xi_t=0\}=P\{\xi_t=1\}=j/2$, тогда детализирующие коэффициенты имеют следующее распределение

$$P \left\{ d_{j,k}^{(\psi)} = \pm (2^{\frac{j}{2}-1} - j \cdot 2^{-\frac{j}{2}}) \right\} = \frac{\sum_{K=0}^i C_{2^{j-1}}^K C_{2^{j-1}}^{i-K}}{2^{2^t}},$$

$$i = \overline{0, 2^{j-1}}, \quad j = \overline{1, n}.$$

На основе этой теоремы построен критерий проверки гипотезы H_0 , он использует χ^2 -статистику Пирсона для проверки гипотезы соответствия эмпирического и теоретического распределений вероятностей детализирующих коэффициентов. Проверка проводилась для каждого уровня детализации $j = \overline{1, n}$, гипотеза H_0 принимается, если на всех уровнях принималась нулевая гипотеза χ^2 -критерия.

Методом статистического моделирования были получены оценки вероятностей ошибок первого и второго рода. Оценка вероятности ошибки первого рода составила 0,045 при уровне значимости $\alpha=0,05$. Для оценки вероятности ошибки второго рода β моделировалась случайная последовательность с вероятностью $p=P\{\xi_t=1\}$, которая изменялась от 0 до 0,45. Результаты моделирования приведены в таблице.

p	0	0,05	0,01	0,1	0,2	0,3	0,4	0,45
β	0	0,01	0,02	0,25	0,028	0,032	0,2	0,54

Результаты моделирования показывают эффективность применения вейвлета Хаара в задаче обнаружения неоднородностей в дискретных случайных последовательностях, которые являются основными математическими моделями дискретных сигналов.

1. Chiann, Ch. A wavelet analysis for time series / Ch. Chiann, P. A. Morretin // Journal of Nonparametric Statistics. – 1998. – Vol. 10, № 1. – P. 1–46.
2. Астафьева, Н. М. Вейвлет-анализ: основы теории и примеры применения / Н. М. Астафьева // Успехи физических наук. – 1996. No. 11. – С. 1145–1170.
3. Чуи, К. Введение в вейвлеты / К. Чуи // М.: Мир. – 2001, 285 с.
4. Андрухаев, Х. М. Сборник задач по вероятностей / Х. М. Андрухаев // М.: Просвещение. – 1985, 160 с.

МОДЕЛИ АДАПТИВНОГО ОБУЧЕНИЯ ЧЕЛОВЕКА И МАШИНЫ

Лутковский В. М., Волков А. В.

Кафедра системного анализа и компьютерного моделирования, кафедра генетики, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: lutkovskiv@gmail.com, andrei@cybergizer.com

На основе сравнительного анализа механизмов обучения человека и машины на базе искусственной нейронной сети (ИНС) предлагается подход к развитию модели адаптивного обучения, используемой в дистанционных образовательных системах. В качестве одного из принципов построения адаптивной обучающей системы предложено использовать теорию адаптивного резонанса (АРТ).

ВВЕДЕНИЕ

Актуальность обсуждаемой темы обусловлена перспективами массового применения систем дистанционного обучения [1]. Современные образовательные системы, такие как KNEWTON [2], тестируют текущий уровень обучаемого студента и адаптируют подачу учебных материалов в зависимости от его уровня подготовки и индивидуальных особенностей. При разработке таких систем используется различные модели: обучаемого студента [3], структуры изучаемой предметной области [4] и ряд других. Цель адаптивных образовательных систем – предоставить обучаемым студентам индивидуальные учебные материалы в удобном для них режиме и тем самым повысить эффективность учебного процесса. Мозг студента работает непрерывно, но если во время бодрствования он в реальном времени обрабатывает текущие данные, то во время сна накопленная информация сортируется и переносится в долговременную память. Это накладывает принципиальное ограничение на темп подачи и максимальный объем информации, усваиваемый человеком за один сеанс обучения.

Изучение принципов работы нервной системы человека привело к созданию обучающих машин на базе искусственных нейронных сетей (ИНС) [5]. Алгоритмы обучения ИНС различаются в зависимости от используемых данных и процедуры подстройки синаптических связей между отдельными искусственными нейронами. При этом этапы обучения и применения машины также разделены во времени. Очевидно, что механизмы обучения ИНС значительно проще, чем реальные механизмы обучения человека, но их результатом как в первом, так и втором случае являются синаптические связи между нейронами, установленные на этапе обучения и зафиксированные на этапе решения практических задач. Это позволяет как человеку так и машине корректно реагировать на входные воздействия (т.е. давать правильные ответы на предъявляемые вопросы) после прохождения обучения и тестирования.

Узким местом адаптивных образовательных систем – сложный и субъективный процесс построения модели структуры (далее – доменной структуры) предметной области [6], поэтому авторами предпринята попытка использовать для решения этой проблемы адаптивную резонансную теорию (АРТ) Гроссберга [6] и соответствующую модель ИНС, так как в основе обучения человека и машины лежат общие принципы. Эта модель используется дважды: и как пример построения структуры доменной области теории ИНС, и как подход к выполнению этой процедуры не специалистом, а машиной (иначе говоря, как подход к ее автоматизации).

I. ПРИМЕР ДОМЕННОЙ СТРУКТУРЫ

Для понимания основ теории искусственных нейронных сетей (ИНС) и приобретения навыков моделирования и обучения однонаправленных ИНС, у студента необходимо сформировать категории: А – моделирование нейрона, В – обучение ИНС, С – однослойная ИНС, D – многослойный персептрон, E – дельта-правило, F – обучение многослойной ИНС [6].

Соответствующая доменная структура предметной области с указанием рекомендуемой последовательности изучения указанных категорий приведена на Рис 1.

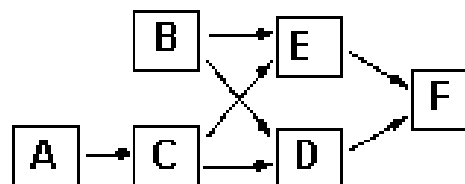


Рис. 1 – Доменная структура

II. ТЕОРИЯ АДАПТИВНОГО РЕЗОНАНСА

Базовые компетенции, сформированные у обучаемого студента на начальном этапе, фиксируются в долговременной памяти и в дальнейшем развиваются при поступлении новых дан-

ных. Так на базе категории «обучение» (А на рис. 1) создаются подкатегории «контролируемое обучение», «конкурентное обучение», «глубокое обучение» и другие. Это приводит к иерархическому усложнению и разветвлению первоначальной упрощенной доменной структуры, представленной выше на рис. 1.

Машины на базе перцептронов, как и человек, могут забывать старый материал, обучаясь на новых данных. Противоречивые требования к механизмам накопления информации и обучения ИНС согласуются в рамках адаптивной резонансной теории С. Гроссберга (S. Grossberg) [6]. ИНС, построенная на базе этой теории (рис. 2), учитывает наличие кратковременной (слой сравнения) и долговременной (конкурентный слой) памяти, а также двунаправленное взаимодействие между ними.

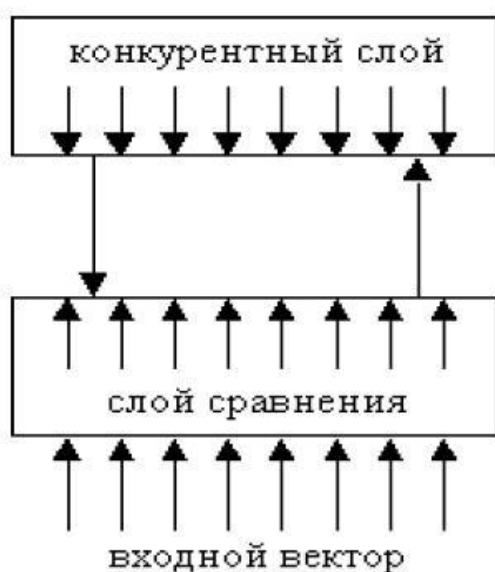


Рис. 2 – ИНС с архитектурой АРТ

Детальный механизм функционирования ИНС этого вида достаточно сложен [6], но в контексте рассматриваемой темы представляет интерес принцип отображения данных, извлекаемых из потока поступающей информации, на долговременную память машины. Новая информация удерживается в кратковременной памяти и сравнивается с категориями, хранящимися в долговременной памяти. При нахождении похожей категории происходит ее уточнение с учетом новой информации, в противном случае в долговременной памяти создается новая кате-

гория. Так называемый параметр бдительности позволяет управлять этим механизмом. Например, при малом значении этого параметра все виды обучения отображаются в долговременной памяти как один большой кластер. Это позволяет адаптировать режим обработки потока входной информации и достичь компромисса между пластичностью и стабильностью памяти машины, что позволяет накапливать новые знания, не разрушая старые.

III. ВЫВОДЫ И РЕКОМЕНДАЦИИ

Сравнение механизмов обучения человека и машины позволяет использовать опыт обучения ИНС и определить возможные направления совершенствования адаптивных обучающих систем. К ним можно отнести следующие:

1. для ускорения обучения и сохранения способности к обобщению объем данных на каждом цикле обучения должен быть ограничен минимумом, достаточным для образования новых категорий в долговременной памяти;
2. при выборе траектории обучения в зависимости от особенностей обучаемого индивидуума метод обучения на примерах на начальном этапе может быть предпочтительнее, чем обучение по правилам.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Трайнев, В.А., Дистанционное обучение и его развитие / В.А. Трайнев, В.Ф. Гуркин, О.В. Трайнев. – М.: Дашков и Ко, 2012. – 294 с.
2. Knewton- [Электронный ресурс] <https://cdn.tclibrary.org/Edlab/Knewton-adaptive-learning-white-paper-1.pdf>
3. Learner Model in Adaptive Learning/ Loc Nguyen et al //The 2008 World Congress on Science, Engineering and Technology (WCSET2008), At Paris, France, Volume: 35. [Электронный ресурс] https://www.researchgate.net/publication/282679719_Learner_Model_in_Adaptive_Learning
4. The Assessment of Knowledge, in Theory and in Practice / Jean-Claude Falmagne, Eric Cosyn, Jean-Paul Doignon, Nicolas Thiery [Электронный ресурс] https://www.aleks.com/about_aleks/Science_Behind_ALEKS.pdf
5. McCulloch. A logical calculus of the ideas immanent in nervous activity/ McCulloch, W, Pitts W. // Bulletin of Mathematical Biology. 1943. Vol. 5. No 4. PP. 115–143.
6. Головкин, В.А. Нейроинтеллект: теория и применение. Книга 2. – Брест Изд. БПИ, 1999. – 264 с.
7. Хайкин С. Нейронные сети. Полный курс 2-е изд. Пер. с англ. – М.: Издательский дом Вильямс, 2006. – 1104 с

СТАТИСТИЧЕСКИЙ АНАЛИЗ ПАНЕЛЬНЫХ ДАННЫХ НА ОСНОВЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ

Меркулов Р. И., Лобач В. И.

Кафедра математического моделирования и анализа данных, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: merkylovecom@mail.ru, lobach@bsu.by

Предлагается подход к статистическому оцениванию параметров моделей панельных данных регрессионного типа, основанный на применении моделей в пространстве состояний и фильтра Калмана. Проведены компьютерные эксперименты, подтверждающие эффективность данного подхода.

ВВЕДЕНИЕ

Панельные данные представляют собой информацию, собранную для разных объектов в последовательные моменты времени. Применение таких наблюдений позволяет специфицировать и оценивать более сложные модели, чем применение одной пространственной выборки или одного временного ряда. Математически панельные данные описываются случайным полем. Примером панельных данных могут служить показатели хозяйственной деятельности совокупности предприятий, которые собираются каждый год. В этом случае получается массив данных, в котором содержатся и данные об однородных объектах за один и тот же период времени, и последовательные значения одной экономической переменной в различные периоды времени. Но если совокупность предприятий из года в год будет различной, то такие данные уже не будут являться панельными.

I. ПАНЕЛЬНЫЕ ДАННЫЕ РЕГРЕССИОННОГО ТИПА

Панельные данные регрессионного типа в отличие от пространственных данных и временных рядов описываются переменными, содержащими два индекса:

$$y_{it} = \alpha + x_{it}\beta + u_{it}, i = 1, \dots, n, t = 1, \dots, T, \quad (1)$$

где i обозначает номер объекта, t — означает момент времени, α — неизвестный скаляр (общее среднее значение), β — $k \times 1$ вектор коэффициентов, x_{it} является i -ым наблюдением k объясняющих переменных. Во многих приложениях панельных данных предполагается, что ошибки наблюдений представляются в виде:

$$u_{it} = \mu_i + \nu_{it},$$

где μ_i обозначает индивидуальный специфичный эффект, ν_{it} — случайные величины. Предположим, что μ_i — последовательность независимых случайных величин с нулевым математическим ожиданием и дисперсией σ_μ^2 , тогда ν_{it} описывается $AR(1)$ временным рядом:

$$\nu_{it} = \rho\nu_{it-1} + \epsilon_{it},$$

где $|\rho| < 1$, ϵ_{it} — последовательность независимых случайных величин с нулевым математическим ожиданием и дисперсией σ_ϵ^2 , $\nu_{i0} \sim N(0, \frac{\sigma_\epsilon^2}{1-\rho^2})$. Проблема состоит в оценивании неизвестных параметров $\alpha, \beta, \sigma_\mu, \sigma_\epsilon, \rho$. Если ввести матрицу $N \times N$

$$C = \begin{bmatrix} (1-\rho^2)^{\frac{1}{2}} & 0 & 0 & \dots & 0 & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -\rho & 1 \end{bmatrix},$$

то с помощью преобразования

$$w^* = (I_N \otimes C)u,$$

где \otimes — кронекерово произведение матриц, получим модель (1) с независимыми ошибками наблюдений, и оценку параметров можно провести методом наименьших квадратов, однако этот процесс требует громоздких вычислений [1].

II. ПАНЕЛЬНЫЕ ДАННЫЕ В ФОРМЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ

Рассмотрим простейшую модель панельных данных:

$$y_i = (1, \dots, 1)^T \beta + \epsilon_i, i = 1, \dots, K, \quad (2)$$

$$y_i \in R^K, \epsilon_i \sim N_K(0, \Sigma_i).$$

Модель панельных данных (2) можно представить в виде следующей модели в пространстве состояний для каждого отдельного объекта (при фиксированном i):

$$\begin{cases} y_t = \beta + \epsilon_t, \\ \epsilon_t = \phi\epsilon_{t-1} + w_t, \\ w_t \sim N(0, \sigma^2), \epsilon_0 \sim N(0, \frac{\sigma^2}{1-\phi^2}). \end{cases} \quad (3)$$

Традиционный подход к проблеме оценивания параметра β сводится к проблеме нелинейной оптимизации. Чтобы избежать этой трудности, можно представить модель (3) в виде:

$$\begin{bmatrix} \epsilon_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} \phi & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} w_t \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} w_t \\ 0 \end{bmatrix} \sim N(0, W), W = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\beta_t = \beta, (\epsilon_0, \beta_0)^T \sim N(0, G), G = \begin{bmatrix} \frac{\sigma^2}{1-\phi^2} & 0 \\ 0 & 0 \end{bmatrix},$$

$$y_t = \epsilon_t, t = 0, \dots, T.$$

Обозначив

$$\begin{bmatrix} \epsilon_t \\ \beta_t \end{bmatrix} = x_t,$$

получим представление в форме моделей в пространстве состояний:

$$\begin{cases} x_t = \begin{bmatrix} \phi & 1 \\ 0 & 1 \end{bmatrix} x_{t-1} + \begin{bmatrix} w_t \\ 0 \end{bmatrix} \\ y_t = [1, 0]x_t, \end{cases} \quad (4)$$

где y_t – наблюдения, x_t – ненаблюдаемые компоненты.

III. КОМПЬЮТЕРНЫЕ ЭКСПЕРИМЕНТЫ

Для оценивания параметра β будем использовать процедуру фильтрации Калмана. Применительно к модели в пространстве состояний (4) рекурсивные уравнения фильтра выглядят следующим образом:

$$\begin{aligned} \hat{\beta}_{t|t} &= \beta_{t|t-1} + K_t \tilde{y}_t, \\ \tilde{y}_t &= y_t - y_{t|\hat{\beta}_{t-1}} = y_t - H_t^T \beta_{t|\hat{\beta}_{t-1}}, \\ K_t &= \Sigma_{t|t-1} H_t (H_t^T \Sigma_{t|t-1} H_t + W)^{-1}, \\ \Sigma_{t|t} &= (I - K_t H_t^T) \Sigma_{t|t-1}, \\ \hat{\beta}_{t+1|t} &= F_t \hat{\beta}_{t|t}, \\ \Sigma_{t+1|t} &= F_t \Sigma_{t|t} F_t^T, \\ y_{t+1|\hat{\beta}_{t+1|t}} &= H_{t+1}^T \beta_{t+1|\hat{\beta}_{t+1|t}}, \end{aligned}$$

где

$$H_t = H = (1, 0), F_t = F = \begin{bmatrix} \phi & 1 \\ 0 & 1 \end{bmatrix}.$$

В нашем случае матрица переходов F_t известна с точностью до параметра, то есть частично неизвестна. Для оценивания параметра ϕ необходимо либо переформулировать модель в пространстве состояний таким образом, чтобы этот параметр стал частью вектора состояния системы x_t , который можно оценить с помощью фильтра Калмана, либо оценивать эту матрицу с использованием функции правдоподобия или EM-алгоритма. В нашем случае задача состоит именно в оценивании параметра β , поэтому для простоты положим, что в данных компьютерных экспериментах матрица переходов процесса известна, то есть параметр ϕ задан явно [2].

Ниже приведены результаты экспериментов при различных значениях параметров ϕ – параметр матрицы переходов F_t , β – свободный член в исходной модели, T – число наблюдений. Эксперименты проведены на модельных данных, порождённых моделью (4) с заданным параметром ϕ .

Таблица 1 – Результаты компьютерных экспериментов

Эксперимент №	ϕ	T	β	$\hat{\beta}$
1	0.1	500	-0.65	-0.599
2	0.5	200	0.3	0.380
3	0.4	200	0.0	-0.050
4	-0.5	100	-0.2	-0.213

Из проведенных экспериментов видно, что оценки параметра β близки к истинным значениям параметра, породившим реализацию случайного процесса в заданных экспериментах.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Harvey, A. C. Forecasting, Structural Time Series Models and the Kalman Filter / A. C. Harvey // Cambridge University Press, Cambridge – 1989. – 227 p.
2. Меркулов, Р.И. Оценивание параметров и прогнозирование экономических временных рядов с пропусками на основе моделей в пространстве состояний / Р. И. Меркулов // Проблемы прогнозирования и государственного регулирования социально-экономического развития: материалы XVII Междунар. науч. конф. – Минск, 2016 г. Т.3. – С. 253–254.

ON THE CLASSIFICATION OF INFORMATION SECURITY THREATS

Mikhailovskaya L. V., Valakhanovich E. V.

Department of Higher Mathematics of the Military Academy of the Republic of Belarus

Minsk, Belarus

E-mail: ludmila_mi@mail.ru

An analysis of the development of modern information technologies demonstrates a substantial increase in the speed of receiving, processing and presentation of information in various spheres of human activity. At the same time, these new opportunities require serious measures aimed to increase information security of an asset. One of the first steps set out in an information protection algorithm is listing and classification of potential threats.

Without due regard to the issue of security of automated information processing systems, consequences of adoption of advanced technologies may be disastrous: along with their benefits, these technologies make it easier to mishandle, destruct or disclose certain data [1]. For that reason, data protection issues become of even more importance. And one of the first steps in developing an information protection algorithm is determining a list of potential threats to information assets and classifying them.

An information security threat is a probable event, action, process or condition that may result in a loss. Today we know a lot of such threats and group them based on various characteristics. Thus, paper [2] considers information security threats targeted at an entity's hardware. Paper [3] lists internal information security threats of a company (insider threats), along with threats from cyberterrorists, as the most urgent. Authors of paper [4] consider use of email and loss of individual information media (laptops, mobile phones, etc.) to be major sources of information security threats. Paper [5] gives the most detailed classification of threat types:

1. Threats caused by force majeure
2. Organization-level threats
3. Threats relating to human error
4. Technical threats
5. Threats arising at a pre-design stage

However, the authors believe that this classification is yet to be finalized. Thus, threats of the second type are caused by misinterpretations in assessing the situation and planning data protection measures, or, in other words, by human errors, while we have a separate category for those (threat type no. 3). Likewise, when considering threat types 4 and 5, we are likely to note that data protection means and projects are designed by people, who are prone to make mistakes. Therefore, we may include them into group 3 – threats relating to human error.

The authors believe that information security threats should be classified by two basic criteria (their nature and security class), depending on the nature of loss they may cause.

We are proposed to classify threats by their nature in two categories: natural threats and human threats (caused by human activity).

A distinctive feature of natural threats is that they are highly unlikely to realize, but when realized, the consequences may be significant.

According to the authors, sources of human threats may be classified in three groups: willful, technology-related and accidental.

Willful threats are related to willful misconduct by penetrators. (Operator may also be a penetrator, if willfully violating the procedure).

Technology-related threats include technical failures and are caused by internal and external reasons. Externally induced threats, including electrical failures, or mains voltage fluctuation, may be indirectly related to organizational deficiencies or unethical practices of staff. Threats induced by internal reasons may be related to deterioration in hardware reliability or faults in their design or software.

Accidental threats are often related to erroneous or insufficient management activities, or lack of competence of staff. Errors of a human operator are considered accidental when caused by fatigue, negligence or lack of experience. The impact of these threats may range from a trivial loss of time due to unavailability of data to a significant loss of information due to violation of both data integrity and confidentiality.

In addition, security threats are classified based on a source of impact depending on the nature of loss they may cause. For this purpose, all threats are considered based on probability of an attack on assets, and feasibility of the attack in specific context of operation of such assets. Information security threats are also appropriate to be grouped by security categories: threats of violation of integrity, confidentiality and availability.

Thus, we propose the following classification of information security threats (Figure 1).

Due to continuous evolution of means and methods of information processing, transfer and protection, the list and classification of information

security threats may be reviewed to include new types of threats and penetration methods.

Therefore, continuous and reliable protection of information in automated data processing systems involves development of information security threats classification as a first step of a procedure aimed to avoid and prevent such threats.

1. DPC/F4.1 Government framework on cyber security – Information Security Management Framework [ISMF], version 3.3, September 2017.

2. Domarev, V. IT Security. Methodology for Creating the System / V. V. Domarev. – Kiev: DiaSowt, 2010. – 673 pages.
3. Malware Threats [Electronic source]. – Moscow, 2010. – Available at: <http://www.pcmag.ru/elearning/course/lesson.php>.
4. Information security threats [Electronic source]. – Moscow, 2011. – Available at: <http://www.infobezpeka.com/publications>.
5. Mehrhoff, M. IT-Grundschutzhandbuch. Standard – Sicherheitsmapnahmen. Bundesamt fur Sicherheit in der Informationstechnik / M.Mehrhoff. – DBUS-Jahrestagung, 2004.

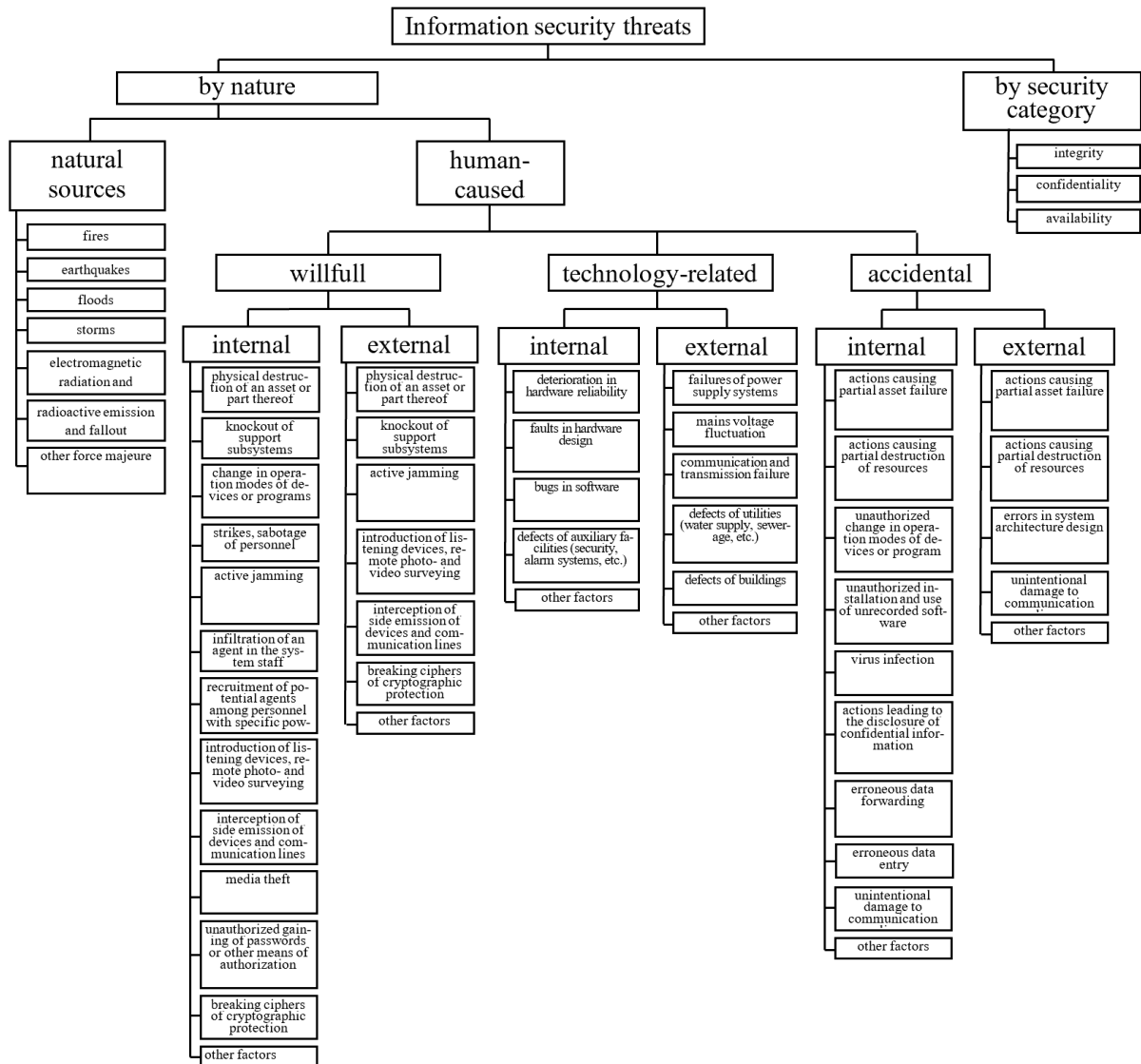


Рис. 1 – Classification of information security threats of an asset

МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ В КОРПОРАТИВНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Моженкова Е. В., Парамонов А. И.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: elena.mozenkova@gmail.com, anton_paramonov@tut.by

Автоматизация процессов сбора и анализа данных корпоративных информационных систем (КИС) в период сопровождения позволит оперативнее локализовать ошибки клиента на стороне разработчика. Эта проблема особенно актуальна в свете разнородности структур баз данных КИС.

ВВЕДЕНИЕ

КИС становятся все более сложными по мере роста и развития отрасли, поэтому поддержание таких систем является актуальной задачей и требует новых подходов. Важным моментом для поддержания КИС является необходимость выполнять работы с более высокими темпами и с меньшими затратами. Разработка функциональной модели формирования общих требований обработки банков данных (БнД) в КИС [1] позволила сформулировать основную цель дальнейшего исследования - построение модели представления информации для различных архитектур корпоративных вычислительных сетей с целью передачи данных на платформу разработчика.

I. АЛГОРИТМ ЛОКАЛИЗАЦИИ ОШИБОК

Предполагается, что процесс локализации ошибки приложения будет включать три этапа:

1. этап подготовки - описание схем и интерфейсов взаимодействия пользовательского приложения с сущностями БД;
2. этап сбора - выборка данных согласно схеме взаимодействия с применением пользовательских фильтров. Фильтр включает действия: выбор пользователем наименования таблицы (далее - менеджер сущностей) и установка значения записи данной таблицы, по которым необходимо осуществлять поиск данных;
3. этап развертывания - имплементация выбранных данных согласно схеме взаимодействия в БД разработчика.

Выполнение общей задачи зависит от всех этапов описанного алгоритма. Каждый последующий этап зависит от результата выполнения предыдущего. Предполагается обработка только реляционных БД, т.к. язык SQL предоставляет возможность реализовать обработку данных путем выполнения специальных запросов [2]. Для правильного выполнения этапа подготовки необходимо формально описать обработку сущностей базы данных КИС. Обработка данных в этом случае является нетривиальной задачей. Для этого, необходимо построить обобщен-

ное представление структуры информации – модель представление информации в КИС.

II. ИНФОРМАЦИОННАЯ МОДЕЛЬ КИС

Большая часть КИС строится на основе трехуровневой архитектуры клиент-сервер [3]. В данной архитектуре можно выделить три логических уровня (см. рис. 1): презентационный уровень (уровень взаимодействия с пользователем), уровень сервера приложений (сосредоточена большая часть бизнес-логики системы), уровень сервера БД (предназначен для хранения состояния системы). В свою очередь, каждый из уровней может делиться на подуровни, например, презентационный уровень может содержать подуровень представления информации в сети Интернет.

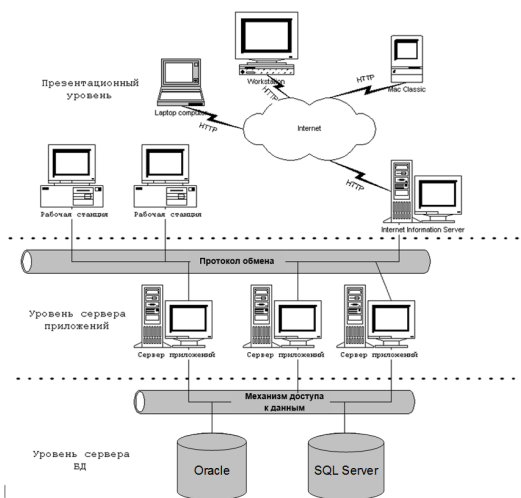


Рис. 1 – Обобщенная структура и основные элементы КИС

Логическая структура организации сервера приложений может быть представлена в виде набора модулей. Под модулем понимается совокупность объектов, их экранных представлений, описаний БД для хранения объектов модуля, выполняющих какую-то общую задачу. Набор модулей, составляющий сервер приложений может изменяться в процессе жизненного цикла КИС.

Каждый модуль может содержать описание схемы БД. Предполагается, что каждый модуль использует собственную схему БД. Реше-

ние о способе организации структуры БД решается на этапе внедрения КИС у заказчика. Обрабатываемый какой-либо модуль приложения, модель представления информации в КИС должна определить соответствующий набор сущностей БД, которые будут обработаны на этапе сбора. Описание сущностей используется в формате метаданных. Среди свойства метаданных прежде всего следует выделить следующие характеристики: метаданные сохраняют системные и прикладные настройки; структура метаданных хранится в формате XML; метаданные представляют собой иерархическую структуру и содержат описание элементов различных типов.

В упрощенном виде модель представления информации, его окружение, временные характеристики и внутреннее устройство можно отобразить схемой взаимодействия, которая представлена на рис. 2.

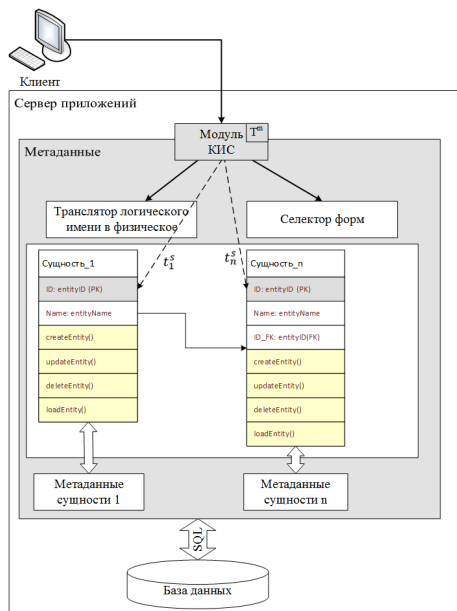


Рис. 2 – Модель представления информации в КИС

В представленной модели вводятся две временные характеристики:

- T^m – расчетный период модуля, т.е. данные для ввода доступны за заданный промежуток времени;
- t_i^S – время актуальности данных при выполнении запроса к сущности i , $i = 1..n$.

Метаданные сущностей, связи между ними, информация, хранимая сущностями выступают исходными данными для локализации дефекта. Источником, ограничивающим набор выбираемых данных, является менеджер сущностей, фильтр установленных значений и временные характеристики. Пользователь сам определяет поле менеджера сущностей, по которому будет осуществляться фильтрация. Программное средство, реализующее процесс локализации ошибки приложения должно учитывать этот набор данных. Формальная модель информационных потоков в разрезе контекста обнаруженно-

го дефекта может быть представлена следующей формулой:

$$M = \langle S_{obj}, T^m, F \rangle,$$

где

- S_{obj} – множество сущностей БД;
- T^m – расчетный период модуля КИС;
- F – множество фильтров менеджера сущностей;

$$S_{obj} = \langle name, m, t_i^S, K_i^S, R_i^S \rangle,$$

где

- name – наименование сущности S_{obj} ;
- m – признак менеджера сущностей;
- t_i^S – время актуальности данных при выполнении запроса к сущности i , $i = 1..n$
- K_i^S – множество кортежей;
- R_i^S – множество связей сущности.

Задача обработки БД заключается в определении набора выбираемых кортежей для каждой сущности, используя связи между таблицами, с учетом заданного пользователем фильтра и временных характеристик.

Процесс локализации дефекта начинается с выбора данных менеджера сущностей по установленному пользователем фильтру. Далее алгоритм должен определить набор связанных сущностей и отобразить необходимые данные. Процесс продолжается до тех пор, пока не будут обработаны все сущности БД. Алгоритм основан на частичной выборке данных из БД, предполагает формирование диапазонов значений фильтруемых данных, соответствующих указанным в исходной формулировке задачи, с последующей проверкой наличия записей в связанных сущностях. Предложенное представление модели позволит формализовать подход к обработке БД КИС с целью повышения эффективности поддержки КИС.

1. Моженкова, Е. В. Модель формирования требований по обработке банков данных в задачах сопровождения / Е. В. Моженкова, А. И. Парамонов // IV Международная научно-практическая конференция (школа-семинар) молодых ученых «Прикладная математика и информатика: современные исследования в области естественных и технических наук», Тольятти, 23-25 апреля 2018 г., Часть № 2. – С. 200 – 203.
2. Singhal, R. Predicting SQL Query Execution Time for Large Data Volume/ R. Singhal, M. Nambiar // ACM SIGKDD, ACM SIGMOD IDEAS '16 Proceedings of the 20th International Database Engineering Applications Symposium: Montreal, QC, Canada — July 11-13, 2016, p. 378-385.
3. Kintonova, A. Zh. Development of Distributed System for Electronic Business Based on Java-Technologies/ A. Zh. Kintonovaa, B. Z. Andassovaa, M. A. Ermaganbetovaa, E. K. Maikibaeva // International Journal of Environmental and Science Education: August 2016, VOL. 11, NO. 10, pp. 3861-3883.

ПОСТРОЕНИЕ ХРАНИЛИЩА ДАННЫХ ДЛЯ ЗАДАЧ ОБРАБОТКИ ДАННЫХ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ

Мысливец О. Р., Рудикова Л. В.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: myslivec.oleg@yandex.ru, rudikowa@gmail.com

В задачах многомерного анализа данных важными моментами являются правильная разработка и построение хранилища данных. В данной статье описывается процесс построения хранилища данных для задач обработки данных пользователей социальных сетей для последующего выявления закономерностей в образовании социальных групп в сети Интернет.

ВВЕДЕНИЕ

Обработка данных пользователей социальных сетей является в настоящий момент довольно перспективным направлением в анализе данных. Анализ поведения пользователей, сбор и накопление данных, анализ характерных признаков пользователей, анализ поведения пользователей, развитие систем рекомендации на основе предпочтений пользователей и прогнозирование связей в социальных группах - наиболее распространенные приложения для анализа данных социальных сетей и интернет-ресурсов. В связи с этим возрастает необходимость в развитии концепций сбора, хранения и обработки информации подобного рода. Для решения задач подобного рода разумно использовать многомерный анализ данных [1]. Несмотря на то, что огромное количество информации накапливается во всем мире, часто подобные данные имеют весьма разрозненный вид и далеки от предоставления общей картины жизни общества.

I. ДАННЫЕ СОЦИАЛЬНЫХ СЕТЕЙ

Данные социальных сетей, которые будут в последствии поступать в хранилище данных собираются с помощью API соответствующих социальных сетей. Несмотря на то, что часть данных многих пользователей может быть скрыта, имеется возможность получить такую основную информацию как [2]:

- базовая информация о пользователе (имя, фамилия, год рождения, город и т.д.);
- информация о месте учебы (школа, университет, институт);
- информация о подписках пользователей;
- информация о постах пользователей.

Информация о постах пользователей является наиболее ценной для анализа, так как, как правило, несет в себе краткую характеристику пользователя, что позволяет охарактеризовать его в некой мере и выделить группы, к которым пользователь может принадлежать. Также каждый пост пользователя может включать геоло-

кацию, хеш-теги и некоторые полезные сведения (текст, видео и аудио файлы). Исходя из всех полученных сведений с помощью создаваемого хранилища станет возможным анализировать пользователей исходя из их предпочтений и выявлять скрытые зависимости и тенденции интересов пользователей.

II. ПРОЕКТИРОВАНИЕ ХРАНИЛИЩА ДАННЫХ

Основным аспектом в концепции проектирования системы на уровне хранения и работы с данными является подход с использованием хранилища данных (Data - Warehouse) – предметно-ориентированной информационной базы данных, построенной на основе схемы «созвездие фактов», специально разработанной и предназначенной для подготовки отчетов и бизнес-анализа с целью поддержки принятия решений. Данные, которые поступают в хранилище данных, доступны, в основном, только для чтения и поступают из внешних источников [3].

Предполагается, что данные будут поступать из различных социальных сетей, а это значит, что для получения и обработки данных необходимо использовать ETL-процесс. Данные социальных сетей являются слабоструктурированными и тот факт, что каждая социальная сеть хранит информацию в своих собственных структурах и предоставляет доступ к ней различными методами, делает процесс получения и очистки данных важным этапом при проектировании всей системы сбора и анализа данных.

Данные из различных социальных сетей собираются путем использования соответствующих API-сервисов. Полученная информация проходит процесс очистки - убираются все неинформативные данные. Далее производится структурирование данных - информация из различных социальных сетей объединяются и записываются в промежуточную базу данных.

Концептуальная модель базы данных для промежуточного хранения информации о пользователях представлена на рисунке 1.

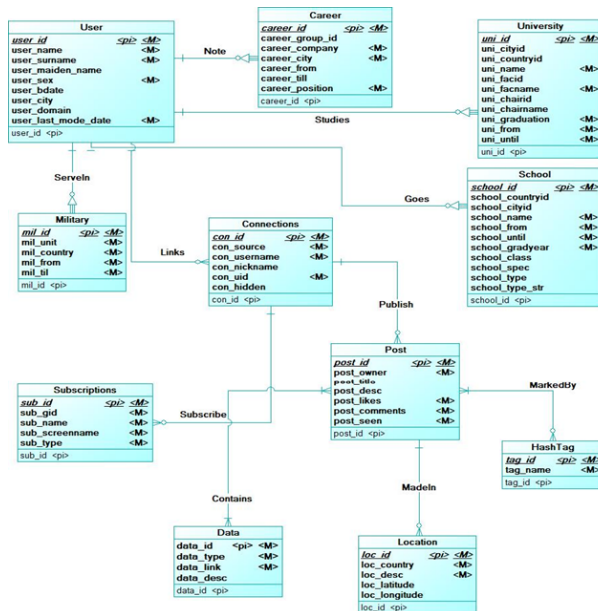


Рис. 1 – Концептуальная модель базы данных для промежуточного хранения информации о пользователях

Так как данные в хранилище будут поступать из различных социальных сетей, важно выявить и реализовать соответствия между неструктурированными данными пользователя и данными, которое будут записываться в хранилище. Основные данные о пользователе, которые возможно получить через API социальных сетей представлены в сущности user (см. рис 2):

User		
user_id	<pi> Integer	<M>
user_name	Text	<M>
user_surname	Text	<M>
user_maiden_name	Text	<M>
user_sex	Byte	<M>
user_bdate	Date	
user_dty	Text	
user_domain	Text	
user_last_mode_date	Date & Time	<M>
user_id	<pi>	

Рис. 2 – Получаемая информация о пользователе

Действия пользователей в социальных сетях, в основном, связаны с написанием постов на своих страницах и репостом сообщений из групп или страниц других пользователей. Данная информация и представляет наиболее ценный интерес. Можно получить следующие данные о записях пользователя:

- данные геолокации;
- хеш-теги;
- группа или сообщество, откуда пользователь взял информацию;
- тип информации.

Основными объектами анализа являются географическое расположение пользователя, тип

информации и набор хеш-тегов. Несмотря на тот факт, что из социальных сетей можно собрать гораздо больше дополнительной информации, вышеперечисленные данные будут являться основой для построения OLAP-куба. Куб (см. рис. 3) реализован на основе схемы звезда и позволит анализировать действия пользователей на основе записей на их странице. Как видно из модели хранилища данных, часть информации не будет использована при анализе пользователей.

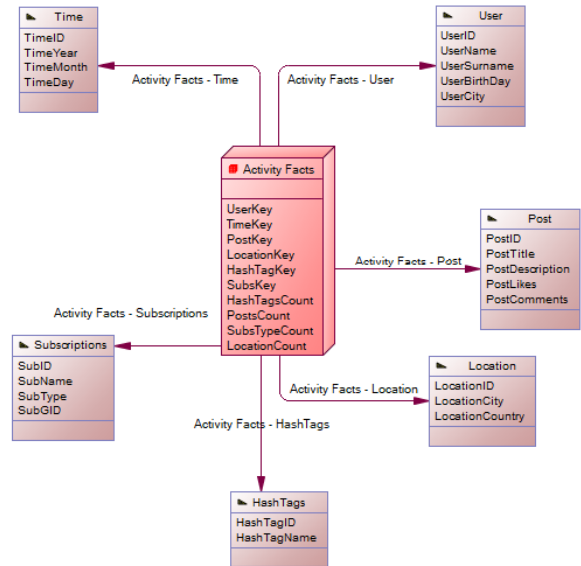


Рис. 3 – Модель OLAP-куба

III. ЗАКЛЮЧЕНИЕ

В данной статье представлена информация о построении хранилища данных для проверки простых гипотез и выявления закономерностей в образовании пользователями социальных групп на основе их предпочтений и записей на их страницах в социальных сетях. Представлена модель OLAP-куба для последующего анализа информации о пользователях. В последствии планируется развитие модели для возможности проверки более сложных гипотез и поиска скрытых предпочтений пользователей на основе имеющейся информации.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Анализ данных социальных сетей: методы и приложения [Электронный ресурс] / ИСП РАН Режим доступа: <http://www.ispras.ru/> – Дата доступа: 18.09.2018.
2. Open API [Электронный ресурс] / Знакомство с API ВКонтакте Режим доступа: <https://vk.com/dev/openapi> – Дата доступа: 18.09.2018.
3. Рудикова, Л.В. Об общей архитектуре универсальной системы хранения и обработки данных практико-ориентированной направленности // Л.В. Рудикова / Системный анализ и прикладная информатика. – Мн.: БНТУ, 2017. – №2. – С. 12-19.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВОЗМУЩЕНИЙ НАПРЯЖЕННОСТИ ЭЛЕКТРИЧЕСКОГО ПОЛЯ, СОЗДАВАЕМОГО ВЕРТОЛЕТОМ

Нефёдов Д. С., Курилович В. В., Дубков А. В.

Кафедра тактики и вооружения ЗРВ факультета ПВО, Учреждение образования "Военная академия Республики Беларусь

Минск, Республика Беларусь

E-mail: VARB_NefedovDS@tut.by

В докладе представлена математическая модель временной структуры возмущений вертикальной составляющей напряженности электрического поля в точке наблюдения на земной поверхности, создаваемых вертолетами на малых и предельно малых высотах. Особенностью модели является учет составляющей электрического заряда вертолета, обусловленной вращением лопастей несущего винта.

ВВЕДЕНИЕ

В процессе решения задач электростатической локации летательных аппаратов (ЛА), основанной на обработке возмущений напряженности электрического поля земли, решаются задачи обнаружения, определения координат и распознавания.

Экспериментальные исследования выявили наличие в амплитудно-частотном спектре (АЧС) сигналов электростатической локации, создаваемых вертолетами, составляющей, обусловленной электрическим зарядом несущего винта (модуляционной составляющей) [1].

Наличие модуляционной составляющей в зарегистрированных АЧС позволяет решать задачу распознавания вертолетов на основании одного из методов, который широко используется для радиолокационного распознавания [2].

Распознавание ЛА по возмущениям естественного электрического поля Земли, по аналогии с задачей радиолокационного распознавания, является статистической задачей. Для ее решения необходимо математическое описание входных воздействий при условии наличия и отсутствия полезных сигналов, создаваемых ЛА распознаваемых классов.

В докладе представлена математическая модель возмущений напряженности электрического поля, создаваемых вертолетами на малых и предельно малых высотах.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ВХОДНЫХ ВОЗДЕЙСТВИЙ

Сигналы в электрической локации описываются математической моделью возмущений напряженности электрического поля, создаваемых электрическим зарядом ЛА (полезный сигнал), и естественными флуктуациями напряженности электрического поля Земли (помеховый сигнал).

Самолеты и крылатые ракеты, в зависимости от высоты полета, представляются точечным или линейно распределенным электрическим зарядом постоянным на интервале наблюдения [3, 4]. Электрический заряд вертолета, ввиду враще-

ния лопастей несущего винта с большой скоростью, включает постоянную q_k и изменяющаяся на интервале наблюдения составляющую электрического заряда q_{vm} , обусловленную вращением несущего винта.

Математическая модель временной структуры возмущений напряженности электрического поля, создаваемых перемещением заряда вертолета и естественными флуктуациями электрического поля Земли в точке наблюдения на земной поверхности может быть представлена выражением:

$$E_{vi} = E_{ki} + E_{vmi} + E_{shi},$$

где E_{ki} – дискретный отсчет напряженности электрического поля, создаваемой зарядом q_k , В/м; E_{vmi} – дискретный отсчет модуляционной составляющей напряженности электрического поля, создаваемой зарядом q_m , В/м; E_{shi} – напряженность флуктуаций электрического поля Земли, В/м; $i = 1...N$ – количество дискретных отсчетов на интервале наблюдения.

Напряженность электрического поля, создаваемая перемещением заряда q_k , описывается известной функцией времени с неизвестными параметрами λ_a, λ_τ [4]:

$$E_{ki} = \frac{\lambda_a \lambda_\tau^3}{(\lambda_\tau^2 + t_i^2)^{\frac{3}{2}}}. \quad (2)$$

Параметры λ_a, λ_τ зависят от величины электрического заряда, высоты, скорости и параметра полета вертолета относительно точки наблюдения.

Напряженность модуляционной составляющей E_{vm} можно аппроксимировать функцией времени $\sin(2\pi f_{vm} t_i)$. Частота модуляционной составляющей f_{vm} зависит от скорости вращения несущего винта вертолета. Полученное эмпирически выражение для данной частоты имеет вид:

$$f_{vm} = \frac{N_{ob}}{k_{vm}}, \quad (3)$$

где N_{ob} – частота вращения вала несущего винта вертолета, об/мин; k_{vm} – постоянный коэффициент равный 12.

Напряженность электрического поля естественных помех E_{shi} является случайным процессом, который в дискретном представлении моделируется отсчетами экспоненциально-коррелированных нормально распределенных случайных величин при помощи формирующего фильтра по рекуррентной формуле [3, 5, 6]:

$$E_{shi} = E_0 + \sum_{k=1}^2 (b_f \sigma_{sh}^2 \text{randn}(x) + a_f E_{shi-k}),$$

где E_0 – математическое ожидание унитарной составляющей электрического поля Земли, В/м; a_f, b_f – коэффициенты формирующего фильтра; σ_{sh} – СКО естественных помех, В; $\text{randn}(x)$ – функция формирования нормально распределенной случайной величины.

Поскольку значения параметров, представленных в выражениях (2) и (3), отличаются для различных вертолетов, будут отличаться и параметры АЧС, создаваемых ими сигналов. В таблице 1, для примера представлены значения параметров сигналов, создаваемых вертолетами Ми-8 и Ми-24 [7], а на рисунках 1 и 2 их АЧС.

Таблица 1 – Параметры сигналов, создаваемых вертолетами

Параметр	Ми-8	Ми-24
λ_a , В/м	0,5...238	0,12...150
λ_τ , с	0,5...4,2	0,5...4,2
N_{ob} , об/мин	192	240
f_{vm} , Гц	16	20

ЗАКЛЮЧЕНИЕ

Представленная математическая модель возмущений напряженности электрического поля, создаваемых вертолетами, может использо-

ваться для решения задач обнаружения и распознавания. Обнаружение целесообразно осуществлять по составляющей АЧС сигналов, создаваемой зарядом q_k , поскольку ее амплитуда на высотах полета вертолета 50...200 метров превышает уровень естественных помех на 20...30 дБ. Для решения задачи распознавания необходимо использовать отличия параметров АЧС сигналов, создаваемых различными типами вертолетов. Так признаком сигнала вертолета является наличие модуляционной составляющей АЧС, а показателем его принадлежности к одному из типов является значение частоты f_{vm} .

1. Нефёдов, Д. С. Экспериментальные исследования сигнала вертолета в электростатической локации / Д. С. Нефёдов, В. В. Курилович, А. В. Дубков // Сборник научных статей курсантов УО "ВАРБ". – 2018. – № 16. – С. 101–106.
2. Небабин, В. Г. Методика и техника радиолокационного распознавания / В. Г. Небабин, В. В. Сергеев. – М.: Радио и связь, 1984. – 153 с.
3. Мелец, А. Ф. Обнаружение и дистанционное измерение электрического заряда маловысотных аэродинамических объектов: дис. ... канд. техн. наук: 20.02.14 / А. Ф. Мелец. – Минск, 1991. – 193 л.
4. Нефёдов, Д. С. Экспериментальные исследования по оценке скорости полета неуправляемых авиационных ракет с использованием модуля пассивной электростатической локации / Д. С. Нефёдов, А. Ф. Мелец // Вестн. Воен. акад. Респ. Беларусь. – 2015. – № 4. – С. 96–103.
5. Нефедов, Д. С. Имитационная модель электростатической многопозиционной системы пассивной локации маловысотных летательных аппаратов / Д. С. Нефедов, А. Ф. Мелец // Информационные технологии в образовании, науке и производстве : II Международная научно-техническая интернет-конференция, 4 декабря 2014 г. Секция: Информационные технологии в производстве и научных исследованиях [Электронный ресурс]. – [Б. и.], 2014.
6. Быков, В. В. Цифровое моделирование в статистической радиотехнике / В. В. Быков. – М.: Советское радио, 1971. – 328 с.
7. Вертолет Ми-24. Техническое описание, кн. IV. Авиационное оборудование. М.: Машиностроение. – 1982 г.

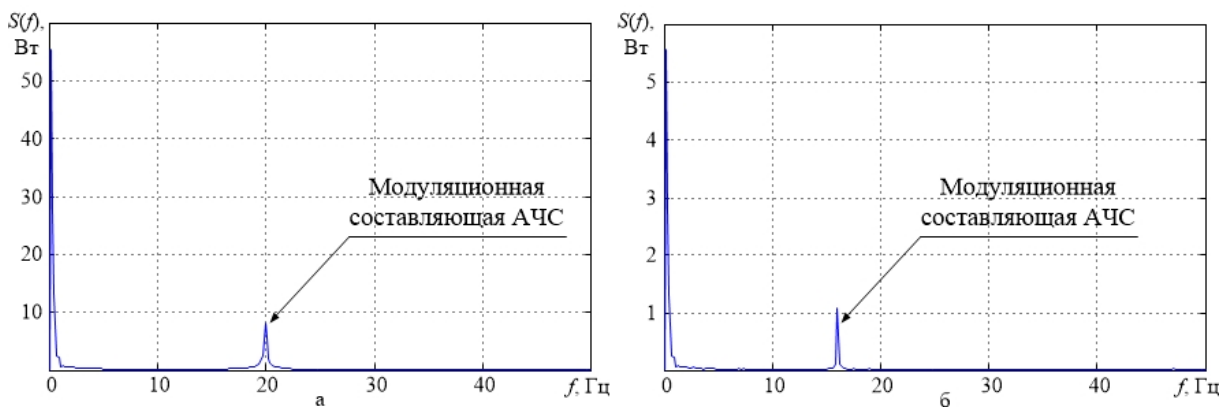


Рис. 1 – Амплитудно-частотные спектры сигналов, создаваемых вертолетами Ми-24 (а) и Ми-8 (б)

О РАЗРАБОТКЕ СИСТЕМЫ ОБРАБОТКИ ДАННЫХ О ПАРФЮМЕРНЫХ БРЕНДАХ

Неживинская А. Ю., Рудикова Л. В.

Кафедра современных технологий программирования, Гродненский государственный университет имени Янки Купалы

Гродно, Республика Беларусь

E-mail: alina.nezhivinskaya@mail.com, rudikowa@gmail.com

Излагаются общие подходы к реализации веб-приложения, которое предназначено для хранения и обработки данных, связанных с парфюмерной продукцией. Приводится основная функциональность программного обеспечения, связанного с разработкой указанного Интернет-ресурса.

ВВЕДЕНИЕ

Покупки в Интернет-магазинах в настоящее время совершаются многими людьми, а количество товаров, которые можно найти на страницах Интернет-сайтов, представлено огромным ассортиментом. С учетом возрастания роли электронной коммерции систематизирование и полная информация необходима как продавцам, так и покупателям. Это особенно заметно на товарах, связанных с парфюмерной индустрией, так как самый главный критерий выбора – запах, который мы не можем ощутить через Интернет. Согласно классической парфюмерной теории все компоненты любого аромата разбиваются на 3 части: верхние, сердечные и базовые. Эти части связаны с процессом испарения аромата. Верхние ноты – это первое, что чувствует человек, вдыхая запах парфюма. Они раскрываются мгновенно и звучат от 5 до 10 минут. Ноты сердца являются ядром аромата, испаряются постепенно и имеют продолжительность звучания от 3 до 5 часов. И, наконец, базовые ноты или шлейф обычно раскрываются спустя некоторое время после сердечных. Такое распределение называется «пирамидой нот» парфюма [1]. Таким образом, пирамида нот может стать своеобразной альтернативой ощущения запаха, ведь, имея в арсенале список знакомых ароматов, из которых состоит парфюм, можно получить общее представление о том, как будет звучать этот парфюм вживую. Поэтому ключевой особенностью разрабатываемой системы является возможность подбора парфюма по любимым ароматным нотам на основе хранящихся в базе пирамид нот парфюмов. Разрабатываемые Интернет-ресурс предполагает создание единой системы, которая позволит автоматически осуществлять сбор данных с сайтов парфюмерных брендов. Полученные сведения будут систематизированы и отображены пользователю в наглядном и удобном формате. Системой также предусмотрен просмотр списка парфюмов, брендов, поиск по параметрам, отображение похожих парфюмов по ароматам и характеристикам. Отображение брендовых магазинов и сайтов также поможет пользователям легко и быстро приобрести

понравившийся продукт. Разрабатываемая система имеет возможность регистрации и оформления подписки, что позволит уведомлять пользователя о новинках.

I. ЭТАПЫ РАЗРАБОТКИ ПРИЛОЖЕНИЯ

Работу над реализацией приложения для хранения и обработки данных, связанных с парфюмерными брендами можно разбить на следующие этапы: анализ существующих решений, таких как Ваш-Аромат.ру и Fragrantica; выявление основных преимуществ и недостатков; проектирование базы данных; получение модели функций разрабатываемого веб-приложения; проектирование пользовательского интерфейса; реализация базы данных и логики взаимодействия с ней; реализация серверной части системы; реализация клиентской части веб-приложения.

II. ОБЩИЕ ПОДХОДЫ К АРХИТЕКТУРЕ РЕАЛИЗАЦИИ ПРИЛОЖЕНИЯ

Представленная система для хранения и обработки данных парфюмерных продуктов представляет собой веб-приложение, включающее клиентскую и серверную части. Основными функциями клиентской части системы являются формирование запросов серверу и отображение полученных данных пользователю. Приложение на стороне сервера отвечает на запросы клиента и выполняет основную бизнес-логику приложения. Система для обработки данных парфюмерной продукции является Интернет-приложением с базой данных, архитектура которого включает следующие уровни (см. рисунок 1): база данных, уровень доступа к данным, уровень бизнес-логики, уровень сервисов, уровень клиента. Уровень клиента представляет собой набор компонентов, отвечающих за предоставление пользователю данных. Реализация всех сервисов, с которыми взаимодействуют компоненты из уровня клиента содержится в уровне сервисов. Далее располагается уровень бизнес-логики, который содержит в себе всю логику, связанную с манипуляцией над данными, реализацию алгоритмов и взаимодействие компонентов между

собой внутри уровня. Для предоставления доступа к данным уровню бизнес-логики необходим уровень доступа к данным, который взаимодействует с базой данных, где в систематизированном виде хранятся вся информация о парфюмерии. Функционирование приложения осуществляется по следующему алгоритму. Клиент, в качестве которого выступает браузер, посылает HTTP-запрос веб-серверу. Далее сервер пересылает этот запрос соответствующему REST-сервису. REST-сервис, в свою очередь, обрабатывает пользовательский запрос и отправляет результат обработки клиенту. По мере необходимости REST-сервис обращается к слою работы с базой данных, непосредственно взаимодействующему с конкретной СУБД.

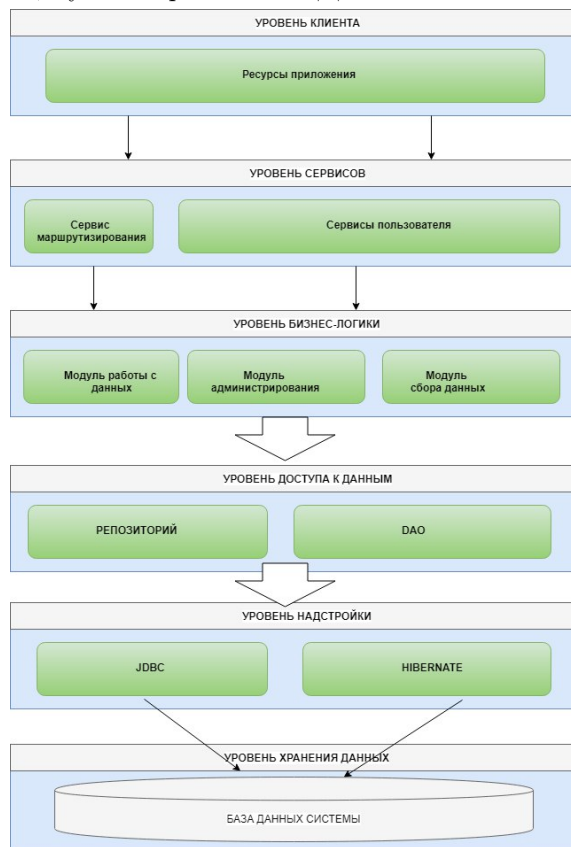


Рис. 1 – Многоуровневая архитектура системы

REST-архитектурный стиль взаимодействия компонентов распределенного приложения, который представляет собой согласованный набор ограничений, учитываемых при проектировании распределенной системы. В определенных случаях использование REST способствует повышению производительности и упрощению архитектуры. Несомненно, обработка запроса приложением сопровождается проверкой наличия у пользователя необходимых прав для совершения запрашиваемой операции. В случае отсутствия этих прав, компоненты страницы или вся страница целиком не отображаются клиенту. Данные для ответа формируются на основе бизнес-логики. Для реализации серверной части

проекта были выбраны: язык программирования Java; СУБД MySQL и технология работы с базой данных Hibernate; технология сборки проекта Maven. Разработка клиентской части (рисунок 2) включает верстку страниц и программирование с помощью Angular 5. Angular представляет собой JavaScript-фреймворк с открытым исходным кодом от компании Google для создания клиентских приложений. Основными целями данного фреймворка являются расширение браузерных приложений, упрощение разработки и тестирования. Прежде всего, он нацелен на одностраничные приложения (Single Page Application), что хорошо подходит для разрабатываемой системы. Приложение, выполняемое непосредственно на стороне клиента в браузере, написано на комбинации из HTML, JavaScript и CSS3. Оно может получать доступ к структурам веб-страницы как к объектам дерева DOM.

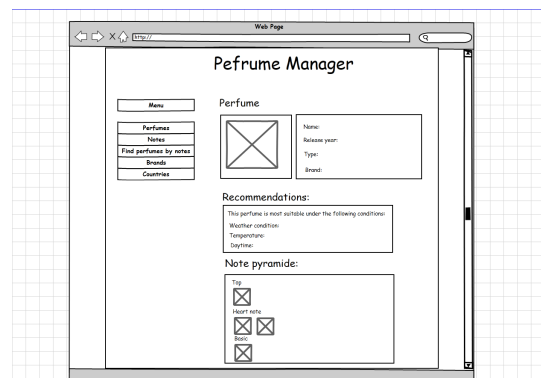


Рис. 2 – Макет страницы информации о парфюме

III. Выводы

Интернет-ресурс с вышеуказанным функционалом будет полезен покупателям парфюмерии, так как использование данного ресурса, позволит сэкономить время на выборе любимого аромата, а так-же быстро и просто получать информацию о новинках в сфере парфюмерии, учитывая уже полюбив-шиеся парфюмерные продукты. Данное приложение также может быть интересно производителям в анализе пользовательских предпочтений. Разрабатываемая система будет расширяемая, а это значит, что в перспективе возможно создание подбора парфюма, учитывая не только стандартные характеристики, но и такие как настроение, погода и цветовая гамма образа.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Неживинская, А. Ю. общие подходы к созданию системы для обработки данных, связанных с парфюмерными продуктам / А. Ю. Неживинская, Л. В. Рудикова // Веб-программирование и Интернет-технологии WebConf 2018 : материалы 3-й Международ. науч.-практ. конф., 12-14 мая 2018 г., Минск. – Минск : Изд. центр БГУ, 2018. – С. 13.

ИНФОРМАЦИОННАЯ СИСТЕМА МОДЕЛИРОВАНИЯ ТЕХНОГЕННЫХ КАТАСТРОФ

Поздняков С. В., Давыдов М. В.

Кафедра информационных технологий автоматизированных систем, кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: sergey.posdnyakov@gmail.com, davydov-mv@bsuir.by

В работе используются методы системного анализа, асинхронного программирования, автоматизированного проектирования информационных систем, технологии хранения данных. Разработанные модели сценариев позволяют определять состав необходимых ликвидационных мероприятий, а также оценивать их эффективность. Кроме того, методы формирования и коррекции планов ликвидации, реализованные в составе информационной системы моделирования, позволят повысить оперативность и эффективность планирования мероприятий при ликвидации крупных техногенных катастроф.

ВВЕДЕНИЕ

Актуальность разработки и моделирования информационной системы по предотвращению и ликвидации последствий чрезвычайных ситуаций обусловлена возрастанием рисков природных и техногенных катастроф. Потребность в разнообразной, своевременной, точной и адекватной информации о состоянии природно-промышленной системы, связанной с предотвращением возможных последствий различных чрезвычайных ситуаций, делает необходимым использование информационных систем, которые отслеживают все возможные состояния природно-промышленной системы, различные влияния на нее, её модели поведения. Рассмотрение вопроса о моделировании такой информационной системы является особо актуальным.

I. АНАЛИЗ СУЩЕСТВУЮЩИХ МАТЕМАТИЧЕСКИХ МЕТОДОВ

Большой раздел современной теории безопасности сложных технических систем посвящен анализу и оценке риска, в которых риск используется в качестве одного из важных критериев безопасности. Риск определяется через функционал F_R вероятности наступления катастрофы (природного или техногенного характера) и величины ущерба:

$$R = F_R(Q, C) \quad (1)$$

При анализе риска вероятность Q и ущерб C можно рассматривать как случайные величины, которые имеют свои распределения $f_Q(q)$ и $f_C(c)$. Ущерб C есть случайная величина, а вероятность события Q также может считаться величиной случайной, потому что она оценивается с некоторой степенью достоверности по ограниченной выборке (например, по статистике аварий). Тогда мы имеем систему двух случайных величин с заданными распределениями и можем ставить вопрос о поиске совместного распределения $f_{QC}(q, c)$. При этом возможно два вари-

анта предположений: случайные величины Q и C независимы (некоррелированы); случайные величины Q и C зависимы (коррелированы).

Таким образом, зная вид законов распределения и значения их параметров $f_Q(q)$ и $f_C(c)$, можно определить вид и значение параметров функций распределения $F_R(r)$ и плотности $f_R(r)$ техногенного риска.

II. АНАЛИЗ И ВЫБОР ЭФФЕКТИВНЫХ АЛГОРИТМОВ И АРХИТЕКТУРЫ

Современные требования к пользовательским интерфейсам веб-ориентированных систем накладывают повышенные требования к интерактивности и удобству использования функций приложения. В связи с этим, язык программирования JavaScript активно используется для работы на стороне клиента, а вместе с ним и весь стек связанных с языком технологий. Из-за специфики веб-приложений, разработка для веб-приложений более трудоёмка, чем написание десктопных приложений. Для того, чтобы решить эту проблему, используют различные веб-фреймворки, основанные на архитектурах MVC и Flux. На данный момент известно несколько хорошо себя зарекомендовавших себя решений: Angular, Ember, Backbone, React. Полноценными фреймворками являются Angular и Ember, Backbone и React-библиотеки для построения пользовательских интерфейсов. Для больших веб-приложений остро встаёт проблема зависимости частей приложения между собой, что влечет значительное возрастание сложности архитектуры веб-системы и падения производительности ее работы.

Одним из путей решения данного вопроса может быть использование связки React.js и архитектуры Flux. Новая, активно развивающаяся и перспективная разработка компании Facebook - React в последнее время привлекает все большее внимание. В сочетании с применением архитектуры Flux библиотеку React.js можно считать полноценным фреймворком. Ос-

новой идеей React.js является компонентно-ориентированный подход, опирающийся на революционную идею сочетать JavaScript-код и разметку компонентов. Flux – это архитектура, базирующаяся на однонаправленном потоке данных и подходе управлением изменением состояния через события.

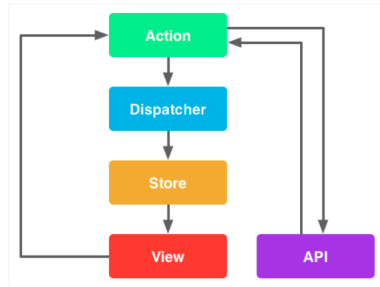


Рис. 1 – Структурная схема архитектуры

Основными элементами Flux-архитектуры являются:

- View – пользовательский интерфейс состоящий из react-компонентов;
- Dispatcher – система регистрации обратных вызовов и реагирования на события;
- Store – хранилище данных на клиенте, которое взаимодействует с View через Dispatcher;
- Action – прослойка между Dispatcher и View для регистрации обработчиков событий действий пользователя.

Типичным подходом в Flux является единственный диспетчер, в котором регистрируются все события приложения и многочисленные хранилища в соответствии с доменными областями приложения.

Таким образом, был рассмотрен подход к решению проблемы возрастания сложности и падения производительности в веб-приложениях, которая связана с особенностью архитектуры платформы, предполагающей единственный диспетчер событий. Подход разделения приложения на зоны ответственности нескольких диспетче-

ров обеспечивает упрощение кода приложения, что сокращает время на расширение и сопровождение кода, а также предотвращает возможную проблему производительности.

III. СТРУКТУРА ПРИЛОЖЕНИЯ

Разработанный пользовательский интерфейс, позволяет работать с основными климатическими характеристиками (температура, влажность, давление, облачность, географические координаты). Была разработана архитектура приложения, произведена сборка проекта (webpack), были созданы независимые друг от друга компоненты.

Главный компонент приложения включает в себя несколько дочерних компонентов, таких, как WeatherDisplay (компонент, включающий получение API из онлайн-сервиса погоды), NavItem (компонент, содержащий список белорусских городов), NavBar (компонент-хедер приложения). Получение API осуществляется в режиме реального времени с сайта <https://openweathermap.org>.

Обработка и получение данных в приложении осуществляется с помощью запроса на указанный выше URL и преобразования их в формат JSON. Ключом для получения данных является zip code (система почтовых индексов города).

IV. СПИСОК ЛИТЕРАТУРЫ

1. Вишняков, А. Д. Общая теория риска / А. Д. Вишняков // Изд. центр «Академия», 2008. – 368 с.
2. Острейковский, В. А. Количественная оценка риска в теории техногенной безопасности сложных динамических систем / В. А. Острейковский // РАН, 2013. – с.12-31.
3. Документация по React.js [Электронный ресурс] – Режим доступа: <https://facebook.github.io/react>.
4. MVC в JavaScript [Электронный ресурс] – Режим доступа: <http://designformasters.info/posts/mvc-javascript>.
5. Flux официальный сайт [Электронный ресурс] – Режим доступа: <https://facebook.github.io/flux/docs>.

ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ

Потараев В. В.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: vic229@rambler.ru

На протяжении последних десятилетий объём текстовой информации, хранимой в информационных системах, резко возрос. Это обуславливает актуальность автоматизированной обработки информации, в том числе текстовой. Нейронные и семантические сети являются эффективными инструментами обработки информации. Рассмотрим применение нейронной сети для решения задачи ответа на вопрос.

ВВЕДЕНИЕ

В современных информационных системах хранятся и обрабатываются довольно большие объёмы данных. Это обуславливает актуальность автоматизированной обработки информации, в том числе текстовой. Для повышения эффективности обработки информации зачастую применяется семантический анализ, то есть учёт смысла данных. Целью данной работы является разработка методов автоматизированной обработки данных, основанных на семантическом анализе.

I. ЗАДАЧА ПОИСКА ОТВЕТА НА ВОПРОС

Информационные системы зачастую используются в качестве инструмента для нахождения ответа на некоторый запрос [1]. Рассмотрим задачу ответа на вопрос, сформулированный на естественном языке.

Для решения данной задачи может быть использован алгоритм, основанный на представлении текста в виде семантической сети. В данной сети узлами являются слова текста, а связями – семантические отношения между словами [2].

II. НЕЙРОННАЯ СЕТЬ ХОПФИЛДА

Одной из моделей представления информации, которые могут быть использованы для классификации данных [3], а также для решения других задач, является модель, основанная на искусственной нейронной сети.

Искусственная нейронная сеть (ИНС) – это математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма [4].

Одним из видов ИНС является нейронная сеть Хопфилда. Основной её особенностью является наличие обратных связей.

Сеть Хопфилда можно определить как динамическую систему с обратной связью, у которой выход одной операции служит входом сле-

дующей операции сети. Каждая операция сети называется итерацией [3]. Схема нейронной сети Хопфилда представлена на рисунке 1.

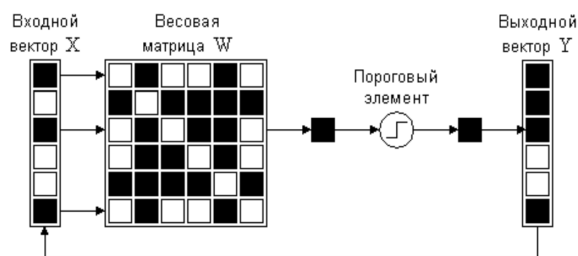


Рис. 1 – Нейронная сеть Хопфилда

К недостаткам сети Хопфилда можно отнести ее сравнительно небольшой объем памяти, то есть в случае записи слишком большого числа образов нейронная сеть перестает их распознавать. Кроме того, ИНС может сойтись к так называемым ложным аттракторам и дать неверный результат [3].

Порядок использования нейронной сети Хопфилда может быть представлен следующим образом:

1. Выбрать параметры сети (размерность входного вектора, множество сохранённых в сети векторов).
2. Рассчитать весовую матрицу на основании параметров.
3. Подать на вход сети распознаваемый входной вектор.
4. Производить перерасчёт выходного вектора сети, пока он не станет постоянным.
5. Полученный вектор является сохранённым сетью вектором, который наиболее «похож» на входной вектор.

III. ПРИМЕНЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ АНАЛИЗА ИНФОРМАЦИИ

Модель семантической сети может быть использована в информационной системе для ответа на вопрос [2].

Модели, основанные на нейронных и семантических сетях, весьма сильно отличаются меж-

ду собой. Так, например, по модели нейронной сети далеко не всегда понятно, о каких объектах идёт речь – даже при небольшом количестве нейронов. Семантическая сеть, напротив, даёт прекрасное смысловое представление о предмете исследования, ее непосредственно можно использовать в качестве результата [5].

Нейронная сеть – это не более, чем сложная алгебраическая функция; для любой нейронной сети можно найти эквивалентную матричную систему уравнений [5].

Тем не менее, модели представления информации, основанные на нейронных и семантических сетях, могут быть использованы для решения одних и тех же задач, например, для решения задачи классификации текстовой информации [3].

Покажем, что модель нейронной сети Хопфилда также может быть использована в информационной системе для ответа на вопрос. Пусть каждому входу сети соответствует некоторое слово (или понятие). Наличие слова в предложении может быть обозначено как наличие сигнала на соответствующем входе. Каждому предложению текста можно поставить в соответствие некоторый вектор.

Порядок использования нейронной сети Хопфилда для ответа на вопрос может быть представлен следующим образом:

1. Выбрать параметры сети (размерность входного вектора – количество уникальных слов текста, множество сохранённых в сети векторов – предложения некоторого текста).
2. Рассчитать весовую матрицу на основании параметров.
3. Подать на вход сети распознаваемое предложение (вопрос).
4. Производить перерасчёт выходного вектора сети, пока он не станет постоянным.
5. Ответ будет содержаться в получившемся выходном векторе.

IV. ПРИМЕР

Предположим, в тексте есть два предложения: «На улице идёт дождь» и «Из-за дождя появились лужи на асфальте». Пример векторов для данных предложений представлен на рисунке 2. Весовая матрица для нейронной сети Хопфилда должна быть составлена с учётом запоминания этих двух предложений.

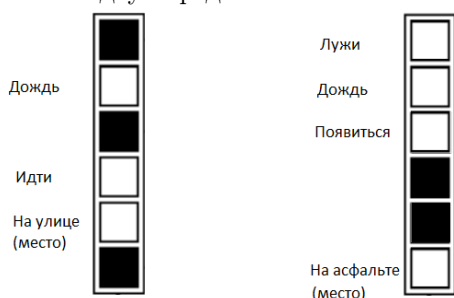


Рис. 2 – Пример векторов для двух предложений

Предположим, нужно ответить на вопрос «Где идёт дождь?». На вход сети подаются 2 сигнала, соответствующие словам «идёт» и «дождь». Очевидно, в результате работы сети на выходе будет вектор, соответствующий словам «дождь», «идти», «на улице» (место). Поскольку «где» – это вопрос определения места, то ответом будет «на улице».

Таким образом, модель нейронной сети может быть использована практически без дополнительных приспособлений для решения задачи ответа на вопрос. Однако, если в вопросе вместо слов текста будут их синонимы, то для получения корректного результата нужно все слова рассматривать в контексте их связей со словами-синонимами. То есть нужен отдельный этап алгоритма, который будет учитывать синонимичность. В случае же семантической сети синонимичность является лишь одним из типов связей, что говорит о большей гибкости семантической сети как модели для обработки текстовых данных.

V. ВЫВОД

Нейронная сеть Хопфилда способна хранить информацию из текста. Её построение может быть автоматизировано.

Нейронная сеть является мощным инструментом распознавания образов. Распознавание предложений текста, в свою очередь, может быть использовано для поиска ответа на вопрос. Семантические и нейронные сети могут использоваться при решении одних и тех же задач. Тем не менее, семантическая сеть является более универсальной моделью для решения различных задач обработки текстовой информации.

1. Информационные системы. [Электронный ресурс]. – Электронные данные. – Режим доступа: http://www.umpro.ru/index.php?page_id=17&art_id_1=261&group_id_4=72. Дата доступа : 23.09.2018.
2. Потараев, В. В. Алгоритм применения семантической сети для поиска ответа на вопрос / В. В. Потараев // Компьютерные системы и сети: Материалы 54-ой научной конференции аспирантов, магистрантов и студентов. – Минск : БГУИР, 2018. – С. 103–105.
3. Серебряная, Л. В. Методы классификации текстовой информации на основе искусственной нейронной и семантической сетей / Л. В. Серебряная, В. В. Потараев // Информатика. – 2016. – №4. – С. 95–103.
4. Рот, В. И. Система распознавания символов с помощью искусственной нейронной сети / В. И. Рот // XIV Международная научно-практическая конференция студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии». – Томск : Изд-во ТПУ, 2016. – С. 54–55.
5. Искусственное сознание (Сети: нейронные или семантические?). [Электронный ресурс]. – Электронные данные. – Режим доступа: <http://alephegg.narod.ru/Survey/NeuralVersusSemantic.htm>. Дата доступа : 23.09.2018.

РАСПОЗНАВАНИЕ ПРОБЛЕМНЫХ СИТУАЦИЙ В СИСТЕМАХ КООРДИНАЦИИ ВЗАИМОДЕЙСТВУЮЩИХ АГЕНТОВ

Ревотюк М. П., Гибулина Е. М., Бруй Н. М.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {rmp, kafitas}@bsuir.by

Предлагается модель и алгоритмы выявления проблемных ситуаций в процессах координации взаимодействующих агентов, ассоциируемые с переходом в недопустимые состояния. Прогнозирование и анализ эволюции состояний процесса координации реализуется на основе построения и анализа интервалов устойчивости решения задачи о динамическом назначении.

I. ПОСТАНОВКА ЗАДАЧИ

Оптимизация управления в системах координации взаимодействующих агентов часто основана на последовательном решении задач о динамическом назначении [1-3]. Такие задачи после линейной свертки локальных критериев сводятся к известным линейным задачам о назначении (ЛЗН) или задачам коммивояжера (ЗК).

Пусть в реальном времени формируется поток открытых ЛЗН

$$\begin{cases} Z_k = \sum_{i=1}^m \sum_{j=1}^n c_{ij}^k x_{ij}^k \rightarrow \min \\ \sum_{i=1}^m x_{ij}^k = 1, j = \overline{1, n}; \sum_{j=1}^n x_{ij}^k = 1, i = \overline{1, m}. \end{cases} \quad (1)$$

Решение (1) есть вектор назначений строк матрицы коэффициентов ее столбцам

$$R_k = \{r_j = i | x_{ij}^k = 1, i = \overline{1, m}, j = \overline{1, n}\}. \quad (2)$$

Известно, что наиболее эффективные для решения задачи (1) алгоритмы венгерского метода [3,4] строятся с учетом особенностей двойственной задачи

$$\begin{cases} Z_k = \sum_{i=1}^m u_i + \sum_{j=1}^n v_j \rightarrow \max \\ c_{ij}^k - u_i - v_j, i = \overline{1, m}, j = \overline{1, n}. \end{cases} \quad (3)$$

При отсутствии возмущений принятому на любом этапе k решению соответствует условие оптимальности паросочетания агентов и назначенных им задач: $c_{ij}^k = u_i + v_j, i = \overline{1, m}, j = \overline{1, n}$. Нарушение такого условия следует рассматривать как проблемную ситуацию, приводящую к изменению (2). Отсюда возникает задача оценки интервалов устойчивости решений открытых и закрытых ЛЗН: для каждого элемента c_{ij}^k в (1) необходимо найти интервал (a_{ij}^k, b_{ij}^k) , в котором значения таких элементов могут быть изменены без нарушения структуры оптимального назначения (2).

Контроль на этапе $k + 0$ условия

$$c_{ij}^{k+0} \notin \{(a_{ij}^k, b_{ij}^k), i = \overline{1, m}, j = \overline{1, n}\} \quad (4)$$

позволяет распознать проблемные ситуации с паросочетаниями агентов, представленными номерами строк $A_k = \{i | c_{ij}^k < \infty, j = \overline{1, n}\}$, и задач, представленными номерами столбцов $T_k = \{j | c_{ij}^k < \infty, i = \overline{1, m}\}$. Элементы $A_k \times T_k$ определяют область возмущения текущего решения, которая должна быть учтена процедурой координации системы агентов [5]. Очевидно, что расширение (4) допускает для групп агентов отражение иерархии вложенности таких областей.

Далее представлены эффективные алгоритмы выделения интервалов устойчивости решения ЛЗН и ЗК, пригодные для распознавания на основе (4) проблем назначения агентам решаемых ими независимых и зависимых задач [2].

II. ПРОГНОЗ РЕШЕНИЯ НЕЗАВИСИМЫХ ЗАДАЧ

Пусть граф оптимального паросочетания агентов и независимых задач представлен элементами $E_m = \{(r_j, j) | (r_j \leq m), j = \overline{1, n}\}$, а оставшиеся элементы матрицы образуют $E_u = \{(i, j), i = \overline{1, m}, j = \overline{1, n}\} \setminus E_m$. Очевидно, что $|E_m| = m$, а $|E_u| = m(n - 1)$, что приходится учитывать при оценке вычислительной сложности алгоритма. Обозначим $\delta_m(x, y)$ и $\delta_u(x, y)$ – допустимое изменение веса любого ребра $x \rightarrow y$, где $(x, y) \in E_m \cup E_u$. Интервалы значений веса ребра, в которых назначение ребра остается неизменным, для ЛЗН вида (1) определяются из элементарных рассуждений:

$$\begin{cases} I_{min}^m(x, y) = (-\infty, c_{xy} + \delta_m(x, y)], (x, y) \in E_m; \\ I_{min}^u(x, y) = [c_{xy} - \delta_u(x, y), +\infty), (x, y) \in E_u. \end{cases}$$

Говорят, что ребро графа оптимального паросочетания скрыто, если его вес увеличен так, что ребро больше не является частью существующего решения. Ребро $x \rightarrow y$ скрывается назначением веса из интервала $(c_{xy} + \delta_m(x, y), +\infty)$.

Для элементов оптимального паросочетания справедливо $c_{xy} = u_x + v_y, (x, y) \in E_m$. Пусть

оценка оптимального решения ЛЗН есть

$$Z^0 = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^0 = \sum_{i=1}^m u_i^0 + \sum_{j=1}^n v_j^0. \quad (5)$$

Нетрудно показать, что если скрыто ребро $x \rightarrow y$, $(x, y) \in E_m$, то процесс реоптимизации, начинающийся в вершине x , завершится в вершине y , потенциал которой не изменится [4]. Меняется только потенциал строки u_x , а в соответствии с (5): $u_x^1 - u_x^0 = Z^1 - Z^0$, $x \in \overline{1, m}$. Здесь нулевой верхний индекс использован для пометки исходного, а единичный – нового решения.

Отсюда следует, что $\delta_m(x, y) = Z^1 - Z^0$. Интервал безопасного изменения значений веса скрываемого ребра определяется выражением

$$\begin{aligned} I_{min}^m(x, y) &= (-\infty, c_{xy} + Z^1 - Z^0] = \\ &= (-\infty, c_{xy} + u_x^1 - u_x^0], (x, y) \in E_m. \end{aligned} \quad (6)$$

Использование разности потенциалов исключает необходимость прямолинейного вычисления оценок решений задачи (1), требующего $m + n$ шагов.

Подобным образом рассматриваются ребра, не принадлежащие оптимальному решению, когда $c_{xy} \geq u_x + v_y$, $(x, y) \in E_u$. Если вес таких ребер менять в интервале $(u_x + v_y, +\infty)$, то структура решения (2) остается неизменной.

Пусть $\epsilon_u(x, y) = c_{xy} - \delta_u(x, y)$. Будем считать, что ребро графа паросочетания открыто, если его вес уменьшен так, что это ребро становится частью нового оптимального решения. Открытие ребра наступает в случае, когда $c_{xy} \leftarrow \epsilon_u(x, y)$. Известен метод определения $\epsilon_u(x, y)$, $(x, y) \in E_u$ на основе построения вспомогательного графа G_a из графа оптимального паросочетания путем удаления всех дуг, инцидентных вершинам x и y . Если выполнить уменьшение веса $c_{xy} \leftarrow \epsilon_u(x, y)$, то для графа G_a , дополненного ребром (x, y) , получим оптимальное паросочетание степени m с оценкой Z^a . Таким образом, $\epsilon_u(x, y) = Z^0 - Z^a$ – нижняя граница интервала устойчивости, поэтому

$$I_{min}^u(x, y) = [Z^0 - Z^a, +\infty), (x, y) \in E_u. \quad (7)$$

Однако определение интервала значений веса ребер на основе (7) не является эффективным. Для каждого из ребер придется решать ЛЗН, размер которых $(m - 1)(n - 1)$.

Предлагается вместо выражения (7) воспользоваться выражением (6), инвертируя направление шагов процесса построения интервала. Конечная граница интервала $I_{min}^m(x, y)$ станет начальной границей интервала $I_{min}^u(x, y)$. Нулевой шаг в (6) становится решением ЛЗН для гарантированно приводящего к открытию ребра значения $c_{xy} = -\infty$, а единичный шаг соответствует решению ЛЗН с исходной матрицей. В результате получаем

$$\begin{aligned} I_{min}^u(x, y) &= (-\infty + Z^0 - Z^1, +\infty] = \\ &= (-\infty + u_x^0 - u_x^1, +\infty], (x, y) \in E_u. \end{aligned} \quad (8)$$

Вычислительная сложность реализации (6) и (8) – $O(n^4)$.

III. ПРОГНОЗ РЕШЕНИЯ ЗАВИСИМЫХ ЗАДАЧ

Известно, что алгоритм оценки устойчивости решения ЗК, определяющего оптимальное паросочетание агентов и зависимых задач, имеет экспоненциальную сложность. Однако результат решения ЗК соответствует решению, например, методом ветвей и границ, закрытой ЛЗН с ограничениями:

$$\begin{cases} Z_k = \sum_{i=1}^n \sum_{j=1}^n c_{ij}^k x_{ij}^k \rightarrow \min \\ \sum_{i=1}^n x_{ij}^k = 1, j = \overline{1, n}; \sum_{j=1}^n x_{ij}^k = 1, i = \overline{1, n} \\ u_i - v_j + n x_{ij}^k \leq (n - 1), i, j = \overline{2, n}, i \neq j. \end{cases} \quad (9)$$

Таким образом, применение (6) к результату решения (9) практически обеспечивает полиномиальную сложность формирования интервалов устойчивости назначения агентам зависимых задач и реакции на незначительные изменения оценок назначения.

Проведенный статистический эксперимент по оценке количества скрытых элементов матрицы ЛЗН, соответствующей оптимальному гамильтонову циклу, показал, что максимальное количество формально скрытых элементов не превышает значения $O(n)$. Если учесть, что общее количество скрытых элементов открытой ЛЗН $n(n - 1)$, то в первом приближении вероятность ошибки в оценке интервала устойчивости не превышает $(n - 1)^{-1}$ (здесь n – размерность матрицы ЗК, $n > 1$).

Таким образом, полученные интервалы устойчивости оптимального паросочетания определяют область возмущения текущего назначения агентов подлежащим решению задачам. Их наличие позволяет усилить логические условия отказа от итераций пересмотра решения при несущественных изменениях параметров внешней среды, что ускоряет реакцию системы управления на обработку информации о состоянии агентов и задач.

1. Spivey, M.Z. The Dynamic Assignment Problem/M.Z. Spivey, W.B. Powell//Transportation Science. –2004. –No. 4. –P. 399–419.
2. Korsah, G. A. A Comprehensive Taxonomy for Multi-Robot Task Allocation/G.A. Korsah, M.B. Dias, A. Stentz [Electronic resource]. – Mode of access: <http://ashesi.org/wp-content/uploads/2016/03/G.-Ayorkor-Korsah.pdf>. – Date of access: 14.01.2018.
3. Toroslu, I.H. Incremental assignment problem/I.H. Toroslu, G. Üçoluk//Information Sciences. – 2007. –Vol.177. –P. 1523–1529.
4. Ревотюк, М. П. Быстрая оценка интервалов устойчивости решения линейных задач о назначении/ М. П. Ревотюк, М. К. Кароли, П. М. Батура//Доклады БГУИР. –2013. – № 5(75). – С. 30–36.
5. Gombolay, M.C. Fast Scheduling of Multi-Robot Teams with Temporospatial Constraints/M.C. Gombolay, R.J. Wilcox, J.A. Shah [Electronic resource]. – Mode of access: <http://roboticsproceedings.org/rss09/p49.pdf>. – Date of access: 24.02.2016.

РЕОПТИМИЗАЦИЯ ПЛАНОВ УПРАВЛЕНИЯ РЕСУРСАМИ В ПОТОКАХ РАБОТ

Ревотюк М. П., Грабовский Д. В.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {rmp, kafitas}@bsuir.by

Предлагается модель и инкрементальный алгоритм оптимального управления ресурсами в потоках работ, когда порядок порождаемых деревьев вариантов существенно меньше порядка графа сети работ. Учет текущего состояния сети снижает сложность пересмотра планов до линейной зависимости от объема сканируемого пространства.

I. ПОСТАНОВКА ЗАДАЧИ

Задачи оптимизации управления потоками работ обычно формулируются в терминах задач о динамическом назначении задач агентам. Агент рассматривается как ресурс, назначаемый для обслуживания заявок. Практически всегда такие задачи сводятся к известным задачам дискретной оптимизации, таким как линейная задача о назначении (ЛЗН) или задача нескольких странствующих коммивояжеров. Однако необходимость учета реальных отношений между агентами и задачами приводит к экспоненциальной сложности алгоритма формирования оптимального назначения. Подобная сложность приводит к задержке момента назначения заданий, снижая эффективность системы агентов. Традиционные приемы использования различного рода аппроксимаций часто неработоспособны из-за недостаточной конкретизации и определенности формируемых решений.

Известно, что классические ЛЗН в виде

$$\begin{cases} Z = \sum_{i=1}^M \sum_{j=1}^N c_{ij} x_{ij} \rightarrow \min \\ \sum_{i=1}^M x_{ij} = 1, i = \overline{1, M}; \sum_{j=1}^N x_{ij} = 1, j = \overline{1, N} \end{cases} \quad (1)$$

характеризуются вычислительной сложностью $O(K^3)$, где $K = \max(M, N)$. Достаточно часто возникает потребность пересчета задачи (1) после изменения ее исходных данных. Например, при учете моментов времени появления работ или исполнителей, начала и окончания работ можно ставить ЛЗН при изменении состояния портфеля заявок[1]. Варианты ЛЗН отличаются лишь изменением некоторых элементов строки матрицы. Итерация расчета для включаемой строки имеет вычислительную сложность $O(K^2)$, что побуждает использовать наследование результатов предшествующего решения путем его реоптимизации[2,3].

Предмет рассмотрения – способы учета наследования решений ЛЗН для ускорения решения взаимосвязанных задач [1] в реальном времени. Без потери общности изложение будем вести для случая матричной постановки ЛЗН, однако

предлагаемый подход применим и для случая ее графовой постановки. Цель работы – расширение схемы инкрементального алгоритма решения ЛЗН [3] на случай решения потока взаимосвязанных задач.

II. МОДЕЛЬ СОСТОЯНИЯ ЗАДАЧИ

Преимущества идеи реоптимизации ЛЗН требуют при ее реализации экономичного способа представления области определения задачи. В случае динамических ЛЗН, определенных в матричной форме, можно выделить следующие операции:

- включение новых строк или столбцов;
- исключение существующих строк или столбцов;
- изменение значений элементов матрицы весовых коэффициентов.

Пусть для хранения матрицы весовых коэффициентов выделена память, соответствующая матрице $C(M, N) = (C_{ij}, i = \overline{1, M}, j = \overline{1, N})$. Матрица текущей ЛЗН $c(m, n) = (c_{ij}, i = \overline{1, m}, j = \overline{1, n})$ является подматрицей матрицы $C(M, N)$, но для работы в реальном времени желательно исключение копирования или бесполезной инициализации.

Операции включения и исключения строк или столбцов очевидным образом реализуются на множествах номеров дуг в списке дуг.

Рассмотрим способ выделения подлежащих реоптимизации строк и столбцов.

Пусть в реальном времени формируется поток задач

$$\begin{cases} Z_k = \sum_{i=1}^m \sum_{j=1}^n c_{ij}^k x_{ij}^k \rightarrow \min \\ \sum_{i=1}^m x_{ij}^k = 1, i = \overline{1, m}; \sum_{j=1}^n x_{ij}^k = 1, j = \overline{1, n}, \end{cases} \quad (2)$$

нумеруемых индексом $k = 1, 2, \dots$. Прямолинейный подход выявления изменившихся строк и/или столбцов на основе поэлементного сканирования матриц с индексами k и $k+1$ характеризуется вычислительной сложностью $O(MN)$, хотя объем фактических изменений будет $O(mn)$. Для отображения изменения строк матрицы будем

использовать вектор

$$X^k(i) = k \cdot (c_{ij}^k \equiv c_{ij}^{k-1}), i \in \overline{1, m}, j \in \overline{1, n}, k > 0,$$

а изменения столбцов отобразим вектором

$$Y^k(i) = k \cdot (c_{ij}^k \equiv c_{ij}^{k-1}), i \in \overline{1, m}, j \in \overline{1, n}, k > 0.$$

Если начальное состояние этих векторов $X^0(i) = 0, i \in \overline{1, m}, Y^0(j) = 0, j \in \overline{1, n}$, а матрица коэффициентов ЛЗН при решении задачи минимизации $c_{ij}^0 = \infty, i \in \overline{1, m}, j \in \overline{1, n}$, то сложность выделения множеств изменения строк и столбцов матриц пропорциональна количеству измененных элементов матрицы. Алгоритм выделения изменений должен формировать стеки индексов изменившихся строк и столбцов.

Алгоритм учета факта изменения $c_{ij}^{k+1} \leftarrow c_{ij}$ в стеке индексов строк $H_x^{k+1}(t)$ и в стеке индексов столбцов $H_y^{k+1}(t)$ на шаге t формирования матрицы с индексами k и $k+1$ имеет вид:

```

 $h_x^{k+1}(0) = 0, h_y^{k+1}(0) = 0$ 
for  $i \in \overline{1, m}, j \in \overline{1, n}$  do
  if  $c_{ij}^{k+1} \neq c_{ij}$  then
    if  $X^{k+1}(i) \neq k+1$  then
       $h_x^{k+1}(t) \leftarrow h_x^{k+1}(t) + 1;$ 
       $H_x^{k+1}(h_x^{k+1}(t)) \leftarrow i;$ 
    end
    if  $Y^{k+1}(j) \neq k+1$  then
       $h_y^{k+1}(t) \leftarrow h_y^{k+1}(t) + 1;$ 
       $H_y^{k+1}(h_y^{k+1}(t)) \leftarrow j;$ 
    end
     $c_{ij} = c_{ij}^{k+1};$ 
  end
end

```

Приведенный алгоритм выполняет однократную фиксацию изменения строки или столбца матрицы, а сложность операции сохранения изменившегося элемента матрицы – $O(1)$.

III. АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ

Стек индексов строк $H_x^k(t)$ и стек индексов столбцов $H_y^k(t)$ определяют отображение матрицы $c(m, n)$ актуальной ЛЗН на память матрицы $C(M, N)$ (при этом $m = |H_x^k(t)|$ и $n = |H_y^k(t)|$). Так как строки и столбцы матрицы ЛЗН формально можно не различать, далее будем полагать выбор варианта отображения, когда $m \leq n$.

Наиболее эффективные для решения задачи (1) алгоритмы венгерского метода используют особенности двойственной задачи

$$\begin{cases} Z = \sum_{i=1}^m u_i + \sum_{j=1}^n v_j \rightarrow \max \\ c_{ij} - u_i - v_j \geq 0, i \in \overline{1, m}, j \in \overline{1, n} \end{cases} \quad (3)$$

Здесь неизвестными являются потенциалы строк $\{u_i, i \in \overline{1, m}\}$ и столбцов $\{v_j, j \in \overline{1, n}\}$. Значения потенциалов особого интереса не представляют, но определяют решение задачи (2). Отображение решения (3) будем осуществлять на упорядоченный вариант вектора назначений строк

столбцам

$$R_y(j) = \{i \mid c_{ij} - u_i - v_j \geq 0, i \in \overline{1, m}, j \in \overline{1, n}\}.$$

Схемы известных версий алгоритмов венгерского метода [1-3] совпадают, включая быстрый этап инициализации для формирования начального назначения строк и итерационного дополнения решения для оставшихся строк. Вычислительная сложность этапа инициализации – $O(K^2)$. На этом этапе пытаются выполнить назначение строк, используя операцию приведения матрицы задачи. Приведение состоит в вычитании из элементов столбцов минимальных элементов столбцов. Однако этап инициализации можно исключить, совмещая этап начального назначения строк с этапом последовательного поиска решения для всех оставшихся строк. Такой прием является ключевым для построения инкрементального алгоритма реоптимизации [3,4].

Симметричная структура данных модели исключает необходимость транспонирования матриц для соблюдения условия $m \leq n$. Обратное отображение решения (3) на упорядоченный вариант вектора назначений столбцов на строки

$$R_x(i) = \{j \mid c_{ij} - u_i - v_j \geq 0, i \in \overline{1, m}, j \in \overline{1, n}\}$$

формируется явно либо может поэлементно формироваться неявно любым алгоритмом венгерского метода. Отсюда следует, что векторы $R_x(\cdot), R(\cdot), u$ и v могут размещаться в предварительно выделенных массивах, размерность которых $K = \max(M, N)$. Выбор варианта отображения реализуется проверкой условия $m \leq n$. Переключение между вариантами элементарно реализуется одношаговым изменением указателей на соответствующие массивы.

ЗАКЛЮЧЕНИЕ

Таким образом, предложенный прием поиска оптимального паросочетания работ и исполнителей позволяет исключить холостые шаги инициализации переменных состояния или повторения поиска. В результате вычислительная сложность реоптимизации решений ЛЗН линейно зависит от количества измененных кортежей отношения работ и исполнителей.

СПИСОК ЛИТЕРАТУРЫ

1. Spivey, M.Z. The Dynamic Assignment Problem/M.Z. Spivey, W.B. Powell//Transportation Science. –2004. –№. 4. –P. 399–419.
2. Toroslu, I.H. Incremental assignment problem/I.H. Toroslu, G. Üçoluk//Information Sciences. – 2007. –Vol.177. –P. 1523–1529.
3. Ревотюк, М. П. Реоптимизация решения задач о назначении /М. П. Ревотюк, М. П. Батура, А. М. Полоневич //Доклады БГУИР. –2011. – № 1(55). – С. 55–62.
4. Ревотюк, М. П. Быстрая оценка интервалов устойчивости решения линейных задач о назначении/ М. П. Ревотюк, М. К. Кароли, П. М. Батура//Доклады БГУИР. –2013. – № 5(75). – С. 30–36.

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ-КЛАССИФИКАТОРОВ ДЛЯ ОБНАРУЖЕНИЯ ИЗМЕНЕНИЙ НА СПУТНИКОВЫХ СНИМКАХ

Рогачёв Н. Е.

Механико-математический факультет, Белорусский государственный университет

Минск, Республика Беларусь

E-mail: mikalai.rahachou@gmail.com

В рамках исследования был сформирован массив оптических изображений земной поверхности до и после стихийного бедствия. Для обнаружения обрушенных зданий на этих изображениях был разработан алгоритм обнаружения изменений по результатам сегментации изображений нейронной сетью, основанной на распространенной архитектуре для общей задачи классификации изображений. Результаты работы алгоритма позволяют сделать вывод о высокой эффективности подхода по модификации и дообучению сети-классификатора для интеллектуального анализа изображений дистанционного зондирования.

ВВЕДЕНИЕ

Информация об изменениях на земной поверхности имеет крайне важное значение в процессах мониторинга материальных ресурсов. Для высокой точности обнаружения изменений необходимо выявление сложных пространственных связей на изображении. Важное значение приобретают интеллектуальные методы анализа изображений и, в частности, нейросетевые методы, позволяющие автоматизировать выявление таких связей. В рамках настоящего исследования был разработан алгоритм обнаружения изменений и оценена его работа в условиях, максимально приближенных к условиям реальной прикладной задачи.

I. ИСПОЛЬЗОВАННЫЕ ДАННЫЕ

Для обучения нейронной сети использовался массив данных Inria Aerial Image Labeling Dataset [1]. В общей сложности он содержит 180 снимков размером 5000×5000 пикселей с пространственным разрешением 30 см на пиксель. Изображения разнообразны по плотности застройки, размеру зданий и цветовой гамме, поэтому можно ожидать устойчивости обученной модели к этим факторам.

Из каждого изображения извлекались фрагменты размером 500×500 пикселей с шагом в 250 пикселей, для каждого из фрагментов создавалось 8 копий с различной ориентацией объектов в пространстве. В результате было получено 519840 фрагментов. Пятая часть из них была случайным образом отобрана для валидации обучения.

В качестве конкретного приложения для системы было предложено обнаружение обрушения зданий в результате землетрясения в Гаити в январе 2010 года [2]. Размер снимков составляет, как правило, 19584×19584 пикселей, пространственное разрешение – около 48 см на пиксель.

Из перечня предоставленных снимков была отобрана пара разновременных снимков с пересекающейся областью захвата. После этого на изображении был произведен визуальный поиск разрушенных зданий. Фрагменты изображения, содержащие разрушенные здания, составили новый массив данных для проверки качества работы системы обнаружения.

II. ОПИСАНИЕ АЛГОРИТМА

Предлагаемый метод распознавания принадлежит к группе методов выявления изменений после классификации. Для сегментации изображений в нем была использована сверточная нейронная сеть, на вход которой подаются изображения земной поверхности в видимом диапазоне.

В качестве архитектуры нейронной сети для сегментации было выбрано семейство SegNet [3]. Основная идея архитектуры состоит в формировании «бутылочного горлышка» между сетью-кодировщиком и сетью-декодировщиком, в котором размерность карт свойств уменьшается, а количество карт для каждого сверточного слоя растет (см. рис. 1). Таким образом, сеть вынуждена выделять наиболее информативные высокоуровневые свойства.

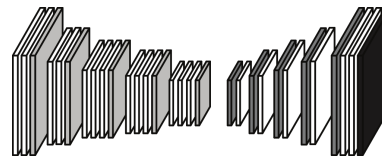


Рис. 1 – Схема сети SegNet

Сеть-кодировщик представлена известной сетью для классификации изображений VGG13. Эти сети обладают проверенными возможностями по выделению высокоуровневых свойств изображений [4].

Недостатком многократного применения пулинга и обратной ему операции является достаточно грубая карта сегментации. Для устранения этого недостатка была использована мо-

дель фильтрации Guided Filter [5]. В её основе лежит предположение о том, что карта сегментации локально линейно зависит от направляющего изображения. Это ограничение не сказывается на качестве фильтрации, но позволяет производить её за линейное время относительно размера изображения.

Итоговая схема алгоритма выглядит следующим образом (см. рис. 2):

- для пары одновременных изображений вычисляются карты вероятности принадлежности к классу «здание»;
- вычисляется карта разности;
- карта разности фильтруется с использованием более раннего изображения в качестве направляющего;
- карта отсекается по положительному граничному значению.



Рис. 2 – Шаги алгоритма

III. РЕЗУЛЬТАТЫ РАБОТЫ АЛГОРИТМА

В качестве начального приближения для кодировщика были использованы веса VGG13, натренированные для задачи классификации изображений из массива ImageNet [6]. Веса декодировщика были инициализированы из равномерного случайного распределения, предложенного в [7]. В качестве функции потерь использовалась функция категориальной перекрестной энтропии, обучение производилось методом градиентного спуска, темп обучения вычислялся с применением алгоритма ADADELTA. Мини-батчи состояли из 2 изображений, обучение производилось в течение 2 эпох. Общее время обучения составило около четырех суток. Значение метрики Intersection-over-Union без применения фильтрации составило 0,549. Применение Guided Filter с параметрами $r = 8$, $\epsilon = 6$ позволило поднять это значение до 0,564 и существенно повысить детализацию и четкость сегментации.

Алгоритм корректно обнаруживает обрушение зданий на фрагментах снимка (см. рис. 3). В отдельных случаях замечены ложноположительные срабатывания в случае оптических искажений и у краев фрагмента. С учетом значительных отличий в условиях съемки между обучающим множеством и данными прикладной задачи, результаты следует считать высокими.



Рис. 3 – Пример обнаружения обрушений

По результатам исследования был сделан вывод о том, что использование нейронных сетей-классификаторов для выделения свойств является эффективной стратегией разработки систем интеллектуального анализа изображений. На примере задачи выявления обрушений было показано, во-первых, что веса нейронных сетей, полученные в рамках общей задачи классификации изображений, представляют собой эффективное приближение для задачи сегментации изображений дистанционного зондирования, а во-вторых, то, что сеть, обученная на изображениях одних регионов, обладает достаточной обобщающей способностью для выявления изменений на изображениях других регионов даже при значительных отличиях в условиях съемки.

При этом разумно предположить, что повышение разнообразия данных в обучающем множестве способно обеспечить еще более высокие результаты. Также к улучшению результатов может привести использование нейросетевых архитектур со сквозными связями. Эти предположения обуславливают необходимость дальнейшей работы над проблемами настоящего исследования.

1. Maggiori, E. et al. Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark / E. Maggiori et al. // IEEE International Symposium on Geoscience and Remote Sensing (IGARSS). – IEEE, 2017. – P. 3226–3229.
2. Open Data Program [Electronic resource] / DigitalGlobe, 2018. – Mode of access: <http://www.digitalglobe.com/opendata>. – Date of access: 01.09.2018.
3. Badrinarayanan, V, Kendall, A., Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Electronic resource] / V. Badrinarayanan, A. Kendall, R. Cipolla. – 2015. – 14 p. – Mode of access: <https://arxiv.org/pdf/1511.00561.pdf>. – Date of access: 01.09.2018.
4. Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Electronic resource] / K. Simonyan, A. Zisserman // 2014. – Mode of access: <https://arxiv.org/pdf/1409.1556.pdf>. – Date of access: 01.09.2018.
5. He, K., Sun, J., Tang, X. Guided Image Filtering / K. He, J. Sun, X. Tang // IEEE Transactions on Pattern Analysis and Machine Intelligence. – IEEE, 2013. – Vol. 35, № 6. – P. 1397–1409.
6. Release VGG16, VGG19, and ResNet50 · fchollet/deep-learning-models [Electronic resource] / F. Chollet. – 2016. – Mode of access: <https://github.com/fchollet/deep-learning-models/releases/tag/v0.1>. – Date of access: 01.09.2018.
7. Glorot, X., Bengio, Y. Understanding the difficulty of training deep feedforward neural networks / X. Glorot, Y. Bengio // Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10). – Society for Artificial Intelligence and Statistics, 2010. – P. 249–256.

О РАЗРАБОТКЕ СИСТЕМЫ ХРАНЕНИЯ И ОБРАБОТКИ ДАННЫХ ПРОИЗВЕДЕНИЙ ХУДОЖЕСТВЕННОЙ ЦЕННОСТИ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ СКЛАДИРОВАНИЯ ДАННЫХ

Рудикова Л. В., Жавнерко Е. В.

Кафедра современных технологий программирования, Гродненский государственный университет имени Янки Купалы

Гродно, Республика Беларусь

E-mail: {rudikowa, eugene.zhavnerko}@gmail.com

В работе излагаются общие принципы построения универсальной системы хранения и обработки данных произведений художественной ценности на основе технологии складирования данных. С использованием структурной и объектно-ориентированной методологий разработаны необходимые модели предлагаемой системы. Реализованы отдельные модули и сделаны выводы об актуальности использования разработки и перспективах развития проекта.

ВВЕДЕНИЕ

В настоящее время развитие методов записи и хранения данных привели к огромному росту объемов, практически, необработанной информации. Конечно, при рассмотрении тех или иных аспектов предметных областей можно указать ресурсы и средства, которые используются для анализа накопленной информации. Однако, целый ряд направлений деятельности различных структур общества требуют построения концепции, а, в дальнейшем, разработки и использования соответствующих систем накопления, расширенного поиска и анализа больших объемов данных. Развитие подходов и концепции к построению систем, связанных с накоплением данных в хранилище и последующим использованием алгоритмов Data Mining является особо перспективным, т.к. позволяет, прежде всего, решать конкретные научно-исследовательские, промышленные, социально-культурные, а также бизнес-задачи. В аспекте разработки специализированного программного обеспечения историко-культурный и научных интерес представляет разработка универсальной интегрированной Интернет-системы, предназначенной для сбора, хранения и обработки данных историко-художественного наследия.

I. О ПРЕДМЕТНОЙ ОБЛАСТИ УНИВЕРСАЛЬНОЙ СИСТЕМЫ ХРАНЕНИЯ И ОБРАБОТКИ ДАННЫХ ПРОИЗВЕДЕНИЙ ХУДОЖЕСТВЕННОЙ ЦЕННОСТИ

В настоящее время в Интернете достаточно ресурсов, связанных с произведениями искусства. Однако используя их, пользователь может получить ограниченную информацию: как правило, изображение произведения искусства с информацией по автору и картине. Проанализировав существующие решения можно сделать вывод, что все они имеют ряд недостатков: отсутствие возможности накопления результатов проведенных исследований, отсутствие обмена опытом, недостаточный функционал вспомогатель-

ного программного обеспечения, баз данных и инструментов анализа накопленных данных, образцов и энциклопедических сведений. Создание соответствующего ресурса позволит хранить расширенные данные об объекте исследования, его характеристиках, материалах, использованных в процессе его создания и т.д., а также осуществлять быстрый и направленный поиск в базе накопленных экспертиз и энциклопедических знаний, формирование необходимых аналитических и исследовательских отчетов. Основная концепция предлагаемой системы для хранения и анализа данных художественной ценности основана на технологии складирования данных. Разработка хранилища данных предполагается с учетом того, что в конечном итоге универсальная система будет предоставлять большой комплекс услуг соответствующим группам пользователей. Очевидно, что ресурсоемкость системы будет расти по мере того как система будет наполняться данными и обслуживать все большее количество пользователей. Немаловажную роль в плане определения концепции построения системы также играет необходимость осуществления аналитической обработки поступающих данных, поиска информации, формирование и хранение итоговых документов, а также предоставление программного интерфейса для доступа к функциональности системы извне. Итак, основным аспектом в концепции проектирования универсальной системы на уровне хранения и работы с данными является подход с использованием хранилища данных, т.е. предметно-ориентированной информационной базы данных, которая специально разработана и предназначена для подготовки отчетов, бизнес-анализа с целью поддержки принятия решений, расширенному интеллектуальному поиску и т.п. по тому или иному направлению. Данные, поступающие в хранилище данных, как правило, доступны только для чтения и характеризуются темпоральным аспектом: информация

загружается в хранилище с определённой периодичностью, поэтому актуальность данных может отставать от транзакционных OLTP-систем. Таким образом, для рассматриваемой предметной области предлагается обобщенная архитектура сбора и анализа данных на базе расширяемого хранилища данных [1]. В данном случае, в качестве расширяемого хранилища данных предлагается подход на основе универсальной платформы хранимых данных, который предоставляет доступ для хранения и последующего анализа данных различной структуры и различных предметных областей, имеющих точки (узлы) стыковки и расширенный функционал с возможностью выбора структуры для хранения данных и последующей внутрисистемной интеграцией. Ядро разрабатываемой системы будет представлять собой хранилище, построенное по схеме «созвездие фактов». Решение в сторону использования схемы «созвездие фактов» обуславливается мощностью платформы БД, и инструментария для реализации запросов. Схема «созвездие фактов» подходит для применения более сложного инструментария для реализации запросов, который в большей степени изолирует пользователей от детальной структуры таблиц, а также для среды с множеством запросов сложной структуры. Основными компонентами архитектура универсальной системы для хранения и обработки данных практико-ориентированной направленности являются: оперативные источники данных; ETL-процесс; хранилище данных; подсистема анализа; пользователи.

II. ПРОЕКТИРОВАНИЕ ВИТРИН ДАННЫХ ДЛЯ ПРЕДМЕТНОЙ ОБЛАСТИ, СВЯЗАННОЙ С ПРОИЗВЕДЕНИЯМИ ХУДОЖЕСТВЕННОЙ ЦЕННОСТИ

Для модели данных предметной области [2], связанной с произведениями художественной ценности и используемой за основу при построении OLTP-системы, использована структурная методология и общие принципы концептуального проектирования. Прежде всего, учтено, что Интернет-ресурс должен предоставить возможность ввода и корректировки информации, связанной с произведениями художественной ценности и авторами работ. В дальнейшем, информация загружается в хранилище, причем, возможно поступление также и из других источников данных, после чего, данные хранилища можно использовать для различных пользовательских аналитических систем. Разработанные витрины данных, которые используются при построении многомерных кубов, позволяют выявить возможные статистические, а также и другие скрытые закономерности с использованием методов DataMining. Кроме того, полученные витрины позволят сформировать наиболее корректную модель типа «созвездие данных» для по-

строения хранилища. Ниже приведен пример витрины данных для стилей произведений художественной ценности (картин). Таблица фактов Стили_Факты содержит атрибуты: Примерное (среднее) время создания картины; Общее количество художественных произведений; Примерная стоимость художественных произведений. Измерениями являются следующие таблицы: Исторический период, Эпоха, Страна, Город, Местоположение, Жанр, Стилль написания, Автор, Школа живописи, Техника написания.

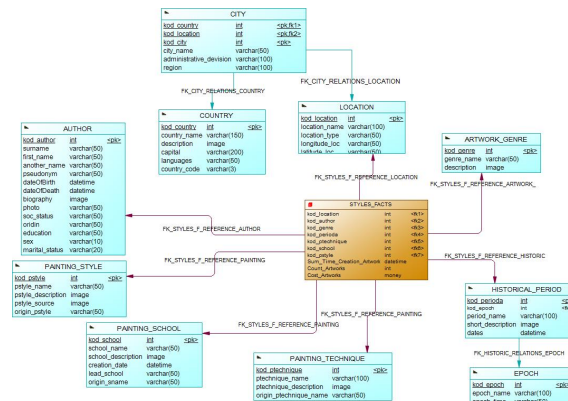


Рис. 1 – Витрина данных для стилей

III. ЗАКЛЮЧЕНИЕ

Таким образом, предлагаемый подход к разработке системы хранения и обработки информации на основе технологии складирования данных, связанной с произведениями художественной ценности, может быть рассмотрена в аспекте создания некоторого хранилища данных, что является актуальной темой исследования. Несомненно, система позволит собрать достаточно обширные сведения по различным аспектам, связанным с произведениями художественной ценности, а также получать необходимые отчеты, проводить структуризацию обработки данных, используя необходимые методы и алгоритмы Data Mining.

Результаты работы получены в процессе выполнения ГПНИ «Разработка методологии и средств построения универсальных систем хранения, обработки и анализа структурированных данных большого объема практико-ориентированной направленности».

IV. СПИСОК ЛИТЕРАТУРЫ

1. Рудикова, Л. В. Об общей архитектуре универсальной системы хранения и обработки данных практико-ориентированной направленности // Л. В. Рудикова / Системный анализ и прикладная информатика. – Мн.: БНТУ, 2017. – №2. – С. 12-19.
2. Рудикова, Л. В. О разработке системы для поддержки экспертизы объектов художественной ценности // Л. В. Рудикова // Технологии информатизации и управления : сб. науч. ст. Вып. 3. В 2 кн. Кн. 2 / под ред. А. М. Кадана, Е. А. Свирского. — Минск РИВШ, 2017. – С. 107-117.

АЛГОРИТМ ФОРМИРОВАНИЯ МЕТРИКИ ЛИЦА В ЗАДАЧЕ ОПРЕДЕЛЕНИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА ПО ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ ЧЕЛОВЕКА

Садов С. В.

Кафедра интеллектуальных систем, Факультет радиофизики и компьютерных технологий Белорусский государственный университет
Минск, Республика Беларусь
E-mail: seregasadov@gmail.com

В работе рассматривается алгоритм формирования метрики лица для случая видеопоследовательности

ВВЕДЕНИЕ

Большинство известных компьютерных систем определения эмоционального состояния человека базируются на получении данных из изображения. Такой подход предполагает наличие достаточного количества информации об анализируемом объекте. Более перспективное направление основано на анализе видеопоследовательности. Предложенный алгоритм позволяет получать всю необходимую информацию об объекте в реальном времени.

I. ОБЩАЯ ХАРАКТЕРИСТИКА ЗАДАЧИ РАСПОЗНАВАНИЯ ВЫРАЖЕНИЯ ЛИЦА ДЛЯ СЛУЧАЯ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

Задача распознавания выражения лица на видеопоследовательности включает в себя следующие основные подзадачи:

- поиск и отслеживание лиц;
- выделение и обработку информативных лицевых характеристик;
- временную видеосегментацию;
- разработку алгоритмов классификации для распознавания выражений лица.

Дадим краткую характеристику компонентам системы распознавания выражения лиц в видеопотоке:

1. Модуль устойчивого поиска и отслеживание лица. Такой модуль является первичной и наиболее важной частью системы. Его главные задачи – принятие решения о нахождении лица в видеопотоке и сужение области обработки в кадрах видеопоследовательности от полного пространства до пространства, содержащего лицо.

2. Модуль видеосегментации. В данной системе под видеосегментацией будем понимать временную, а не пространственную сегментацию каждого видеокadra. Стоит отметить, видеосегментация может выполняться напрямую через классификацию, однако такой подход не всегда дает хорошие результаты. Методы видеосегментации позволяют обнаружить и вычислить характеристики движения лицевых мышц челове-

ка, а также повысить эффективность распознавания изменения мимики лица.

3. Классификатор эмоционального состояния человека, методы которого способны учитывать временные изменения выражения лица. На вход алгоритма поступает набор последовательностей выражений лица (различной длительности) и набор параметров, характеризующих эмоциональные состояния. Задача алгоритма, в соответствии с предложенной информационной моделью описания эмоционального состояния человека – определить вид функции F , которая соотносит каждую последовательность с одним из состояний человека.

II. МОДУЛЬ ПОИСКА И ОТСЛЕЖИВАНИЕ ЛИЦ

Классическая задача нахождения лица, как правило, решается с помощью таких известных методов, как метод главных компонент (Principal component analysis, PCA), линейный дискриминантный анализ (Linear Discriminant Analysis, LDA), искусственные нейронные сети, каскады Хаара [28], метод опорных векторов (Support Vector Machine, SVM), алгоритм AdaBoost (адаптивное усиление) и т. д. Данные подходы к решению задачи нахождения лица на видеопоследовательности обладают рядом недостатков: невысокая точность распознавания при изменениях положения лица перед камерой, высокое число ложных срабатываний, сложность использования алгоритмов в режиме реального времени. Для данного модуля системы распознавания эмоций было решено применить гибридный алгоритм, объединяющий несколько подходов, такие как каскады Хаара и алгоритм AdaBoost, а также использование нейронных сетей. Гибридная схема строится следующим образом: компоненты алгоритма адаптивного усиления принимают решение о наличии лица в кадре видеопоследовательности и отклоняет кадры без лиц. Это позволяет использовать такой подход в реальном времени. Нейронная сеть, получив на вход кадры с лицами формирует окончательное решение о наличии лица в кадре, таким образом выполняя

дополнительную проверку и исключая ложные обнаружения.

III. ВРЕМЕННАЯ ВИДЕОСЕГМЕНТАЦИЯ

Представленный в работе алгоритм видеосегментации не использует предварительное обучение на моделях определенных выражений лица, а основан на непосредственной оценке динамики изменения состояния лицевых мышц, присущую всем эмоциям. Таким образом, предлагаемый алгоритм естественным образом подходит для всех лицевых экспрессий, которые приводят к деформации кожного покрова лица. Важной особенностью предложенного подхода к анализу выражения лица является его направленность на использование временной информации в видеопоследовательностях для улучшения распознавания. Использование предлагаемого подхода позволяет объединить преимущества как геометрических методов моделирования лица, так и методов моделирования, основанных на облике. Предлагаемый блок видеосегментации представляет собой временную сегментацию всей видеопоследовательности на участки, которые содержат кадры одной лицевой экспрессии. Данный алгоритм видеосегментации содержит следующие шаги:

1. Вычисление движений пикселей кадра. Производится вычисление движения пикселей по необходимому количеству кадров видеопоследовательности с помощью метода оптического потока на всей лицевой области.

2. Построение временного ряда. Строится временной ряд, представляющий собой массив данных, характеризующий все значения направлений и скорости сдвигов пикселей. Полученный временной ряд анализируется на наличие сдвигов. Задаются пороговые значения, вычисляются все экстремумы, из которых затем происходит выбор. Выделяются локальные группы пикселей, движения которых соответствуют моментам проявления лицевых экспрессий в видеопоследовательности.

3. Выбор и представление анализируемых лицевых характеристик. Для формализованного описания проявления эмоций в динамике, словесное описание движения их компонент необходимо представить в виде, удобном для восприятия вычислительной системой. Для такого представления подойдет пространственное разбиение ли-

ца на участки, на которых расположены простые лицевые объекты.

4. Применение алгоритма классификации. На данном этапе определяется принадлежность групп смещенных пикселей к лицевым подобластям, полученным на этапе отбора лицевых характеристик. Классификация может осуществляться как с помощью алгоритма AdaBoost с использованием примитивов Хаара, так и с помощью нейронных сетей. Таким образом, зная в какой из подобластей произошло движение пикселей, можно охарактеризовать движение соответствующего лицевого объекта

IV. КЛАССИФИКАЦИЯ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА

После получения подробной характеристики движений лицевых объектов можно переходить к задаче оценки эмоционального состояния человека. Известно, что эмоциональное состояние человека, выражаемое мимикой, можно описать набором состояний лицевых объектов. Массив состояний лицевых объектов, полученный на предыдущем шаге алгоритма, является основой для поиска в нем композиций объектов, описывающих определенное эмоциональное состояние. Такой поиск может осуществляться как с помощью методов нечеткой логики, так и с помощью нейросетей. Такая задача не является однозначной, так как человек не испытывает только одно определенное эмоциональное состояние в период времени. Разработка и описание методов анализа комплексного эмоционального состояния человека является задачей дальнейших исследований.

ЗАКЛЮЧЕНИЕ

Предложен алгоритм формирования метрики лица в задаче определения эмоционального состояния человека для случая видеопотока.

V. СПИСОК ЛИТЕРАТУРЫ

1. 1. Мурыгин К. В. Обнаружение объектов на изображении на основе каскада классификаторов // Искусственный интеллект. 2007. Вып. № 2. С. 104–108
2. 2. Sochman J. AdaBoost with totally corrective updates for fast face detection [Электронный ресурс]. URL: <http://cmp.felk.cvut.cz/sochmj1> (дата обращения: 18.03.2014)
3. 3. Viola P., Jones J. Robust real-time object detection // Proc. of IEEE Workshop on Statistical and Computational Theories of Vision, 2001. P. 324–332.

ОЦЕНКА ВЫЧИСЛИТЕЛЬНОЙ СЛОЖНОСТИ ОБОБЩЕННОГО АЛГОРИТМА КЛАСТЕРИЗАЦИИ

Сасин Е. А., Сидорович А. С.

Кафедра электронных вычислительных машин, Белорусский государственный университет информатики
и радиоэлектроники
Минск, Республика Беларусь
E-mail: a.s.sidorovich@gmail.com

В работе рассмотрен обобщенный алгоритм кластеризации данных. Приведены алгоритмы последовательной и параллельной работы алгоритма. Оценена их временная сложность

ВВЕДЕНИЕ

Развитие методов работы с информацией привело к ее увеличению до колоссальных объемов. Для удобства работы с информацией проектируются и используются базы данных. Но в большинстве своем они таких размеров, что получение и анализ данных из них является проблемой. Часто для решения таких задач используется DataMining. Data Mining — собирательное название, используемое для обозначения совокупности методов обнаружения в данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

I. КЛАСТЕРНЫЙ АНАЛИЗ

Одним из методов DataMining является кластеризация (кластерный анализ). Кластеризация занимается разбиением данных на группы (кластеры) на основе схожести определенных признаков. Методы кластерного анализа позволяют решать следующие задачи:

- проведение классификации объектов с учетом признаков, отражающих сущность объектов;
- проверка выдвигаемых предположений о наличии некоторой структуры в изучаемой совокупности объектов;
- построение новых классификаций для установления наличия связей внутри совокупности;
- сжатие данных – если исходная выборка избыточно большая, можно сократить ее, оставив по одному наиболее типичному представителю от каждого кластера.

Опишем обобщенный алгоритм всех используемых методов кластеризации:

1. Случайным образом выбираются центры кластеров;
2. Вычисляются расстояния от каждого объекта до центра каждого кластера;
3. Объекты причисляются к кластерам;
4. Производится пересчет центров каждого кластера;
5. Если центры не изменились, то работа алгоритма заканчивается, иначе продолжаем

работу алгоритма с вычисления расстояния.

Для вычисления расстояний между центрами кластеров используются различные метрики. Тремя наиболее распространенными метриками расстояний являются классическая Евклидова метрика и метрики Чебышева и Манхеттена.

II. ОЦЕНКА СЛОЖНОСТИ

Особое место при построении систем и алгоритмов является оценка их сложности. Оценка сложности помогает реализовывать более эффективное решение в каждой конкретной ситуации. Сложность алгоритма бывает временная и пространственная. Пространственная – показывает затраты памяти на реализацию алгоритма, а временная описывает количество и время выполнения элементарных операций алгоритма. При работе с большими объемами данных подразумевается использование достаточных физических ресурсов данных, а вот временные затраты на реализацию алгоритмов можно эффективно сокращать (без потери точности). Точные затраты времени рассчитать практически невозможно, т.к. это зависит от физических характеристик вычислительной системы: архитектуры и количества процессоров, особенностей компилятора и многих других. Для оценки временной сложности можно использовать O - нотацию, которая использует математическую функцию $f(n)$, которая зависит от количества операций n . Использование данной нотации позволяет описать характер функции $f(n)$ с изменением n : насколько быстро растет эта функция. Данная нотация использует принцип «худшего случая». Стандартные функции используемые в O - нотации в зависимости от степени их роста:

- остоянные функции, которые с ростом n не изменяются, $O(1)$;
- функции с логарифмической скоростью роста $O(\log_2 n)$;
- функции с линейным ростом $O(n)$;
- функции с линейно-логарифмической скоростью роста $O(n \log_2 n)$;
- функции с квадратичной скоростью роста $O(n^2)$;

- функции экспоненциальной скоростью роста $O(2^n)$;
- функции с факториальной степенью роста $O(n!)$.

Стоит заметить, что сложность алгоритма по данной нотации не обязательно должна принадлежать одной из этих групп.

III. СРАВНЕНИЕ СЛОЖНОСТИ ВЫПОЛНЕНИЯ АЛГОРИТМА

При выполнении алгоритма кластеризации на многоядерных процессорах появляется возможность сокращения временных затрат. Алгоритмы кластеризации получают в качестве входных данных n векторов, каждый размерностью m . Количество разбиений на кластеры предопределено в начале и равно k . Пусть t – количество итераций, необходимых для завершения кластеризации.

Сравнению временной сложности последовательного и параллельного выполнения алгоритмов может проводиться без учета сложности вычислений метрик расстояний и нахождения центров кластеров. Т.к. эти алгоритмы усложняют и параллельную, и последовательную реализацию одинаково.

При обработке больших объемов данных последовательно большую часть времени вычисляются расстояния между кластерами и их центры. Таким образом, последовательно делается n одинаковых операций, которые не затрагивают другие вектора. Обобщенная схема последовательной реализации алгоритма представлена на рисунке 1.



Рис. 1 – Последовательная реализация алгоритма кластеризации

Вычислительная сложность последовательного алгоритма кластеризации $O(nmkt)$

Для организации распараллеливания необходимо выполнить за один такт сразу n операций вычисления центров кластеров, а затем столько же для вычисления расстояний. Обобщенный алгоритм параллельной кластеризации можно представить следующим образом:

1. Генерируется k объектов-центров классов.
2. Производится разбивка на кластеры, путем выяснения какой либо меры близости вектора с центром кластера. Для каждого вектора мера близости вычисляется отдельным процессом, параллельно с остальными.
3. Для каждого кластера выполняется поиск нового центра. Т.к. в каждом кластере данные не связаны с векторами из других кластеров, то и эту операцию можно выполнять одновременно для k разных кластеров.
4. Выполняется проверка, изменился ли центр класса. Если да, то выполняется переход к шагу 2. Если нет, то кластеризация завершена.

Обобщенная схема параллельной реализации алгоритма представлена на рисунке 2.

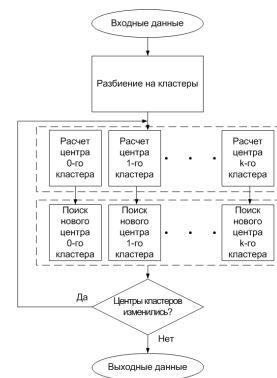


Рис. 2 – Параллельная реализация алгоритма кластеризации

Используя алгоритм нахождения частичной суммы элементов массива путем сдвигания временная сложность вышеприведенного алгоритма составит $O(\log(nmkt))$.

ЗАКЛЮЧЕНИЕ

В ходе выполнения данной работы были получены следующие выводы:

- алгоритмы кластеризации могут быть распараллелены независимо от методов вычисления центров и расстояний.
- алгоритм можно выполнять параллельно, если есть минимум две операции с независимыми данными
- при увеличении количества векторов – производительность параллельных вычислений увеличится

1. Методы и модели анализа данных: OLAP и DataMining. / А. А. Барсегян [и др.]//Спб.: БХВ-Петербург/ –2004.
2. Крупский, В. Н Введение в сложность вычислений. / В. Н. Крупский//–М.: Факториал Пресс. –2006.
3. Arabie, P., Hubert, L. J. Clustering and Classification. / P. Arabie, L. J. Hubert, // Singapore: World Scientific/ –1996

АНАЛИЗ БАЗ ИЗОБРАЖЕНИЙ СЕТЧАТКИ ГЛАЗА ЧЕЛОВЕКА

Сасин Е. А., Сидорович А. С.

Кафедра электронных вычислительных машин, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {sasin, sidorovich}@bsuir.by

Данные тезисы содержат информацию об актуальных базах изображений сетчатки глаза человека, находящихся в открытом доступе. На основании приведенных баз проводятся исследования по классификации и кластеризации изображений глазного дна и их частей. Также проводятся исследования по выявлению сопутствующей информации: кровеносные сосуды и т.д.

ВВЕДЕНИЕ

В настоящее время реалии таковы, что научное сообщество требует проведения и демонстрации результатов исследований на общедоступных базах исходных данных. Это обеспечивает проверяемость и воспроизводимость полученных результатов. Также использование общедоступных баз исходных данных позволяет корректно сравнивать собственные результаты с результатами коллег по всему миру и делать выводы о текущем статусе исследования и намечать дальнейшие пути его развития.

I. ПРОЕКТ STARE

Проект STARE (STructured Analysis of the Retina) был запущен в 1975 году в Калифорнийском университете в Сан-Диего. Автором проекта является Майкл Голдбаум. Изображения и медицинские данные по ним были предоставлены Shiley Eye Center (Калифорнийский университет) и Медицинским Центром Ассоциации Ветеранов [1].

Стандартной процедурой во время визита пациента к офтальмологу является обследование с визуализацией сетчатки. Камера через зрачок глаза может видеть заднюю поверхность глазного яблока и делать ее снимки. На такие снимки попадает зрительный нерв, центральную ямку, сосуды и сетчатку. По этому снимку офтальмолог делает выводы о состоянии здоровья пациента.

Со временем копируются изображения сетчатки и экспертные заключения по ним (диагнозы офтальмологов и результаты консультаций) и объединяются в базы изображений. Такие базы можно использовать для проведения исследований и построения систем автоматической диагностики заболеваний глаза. На вход система получает изображение задней стенки глазного яблока, а на выходе предоставляет диагноз.

От входа до выхода система решает 2 задачи. Первой задачей является обработка изображения сетчатки для выделения областей интереса. Второй задачей является принятие решения о диагнозе на основании информации в областях

интереса в частности и на изображении в целом. Попутно может быть получена дополнительная информация о заболевании. На ее основе может проводиться оценка развития или ремиссии заболевания. Все вышесказанное справедливо и для других баз изображений.

Характеристики базы STARE (см. рис. 1):

- 400 изображений;
- список диагнозов для каждого изображения;
- карта сосудов для каждого изображения;
- 10 изображений, где отдельно выделены артерии и вены;
- 80 изображений, где найден оптический нерв.

С использованием данной базы проводился ряд исследований. Например, поиск сосудов [2] и поиск зрительного нерва [3].



Рис. 1 – Пример изображений из базы STARE

II. БАЗА АНГИОГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ

В данной базе собраны изображения сетчатки глаза человека, полученные методом ангиографии. Данный метод заключается в контрастном выделении сосудов в исследуемой ткани (или образце). В кровь вводится специальное вещество, которое в определенном спектре или при определенных условиях выглядит очень ярко или очень темно на изображении. Следова-

тельно и сосуды буквально «светятся» на изображении, что облегчает сбор информации о них в автоматическом режиме.

Данные изображения получены в Персидской Офтальмологической клинике Исфаханского университета (Feiz Hospital) [4]. В сборе изображений участвовали 70 пациентов с различными стадиями диабетической ретинопатии: 30 нормальных и 40 аномальных стадий. Следовательно, данная база может применяться как обучающая выборка для нейронных сетей и систем принятия решений для «натаскивания» их на диабетические симптомы.

Характеристики базы изображений (см. рис. 2):

- 70 изображений;
- 8-битные изображения;
- расширение 576 на 720;
- изображения не размечены.

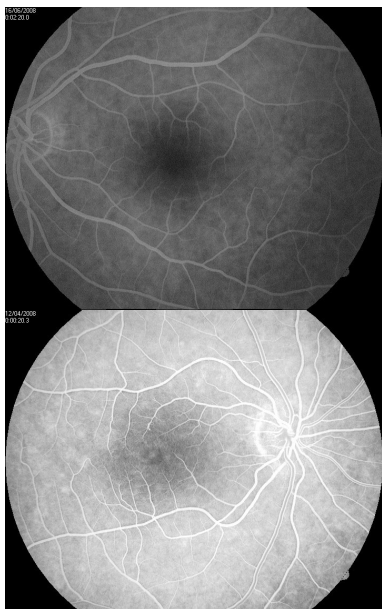


Рис. 2 – Пример изображений, полученных методом ангиографии

С использованием данной базы проводилась работа по распознаванию центральной ямки на задней стенке глазного яблока по сосредоточию сосудов [5].

III. БАЗА ОКТ-ИЗОБРАЖЕНИЙ

Развитие и исследование ОКТ-метода, как неинвазивного прижизненного метода исследования тканей, начались в 1993 году после публикации об ОКТ-изображениях сетчатки глаза человека. ОКТ-метод в настоящее время очень широко используется в офтальмологии.

Следует отметить, что термин ОКТ-изображение является условным, т.к. после исследования образца тканей томограф выдает массив значений размерностью $M \times N$, а представление этих данных в виде изображения осуществляется специальным программным обеспечением.

Работа с ОКТ-изображением подразумевает наличие этапа предобработки, который включает в себя сбор изображения из В-сканов и фильтрация такого изображения от артефактов, которые обусловлены самой ОКТ-технологией.

Приведенная база содержит 50 изображений сетчатки, которые сделаны с помощью цифрового микроскопа и 50 изображений этих же пациентов, которые сделаны методом ОКТ [6]. Пример данных представлен на рисунке 3.

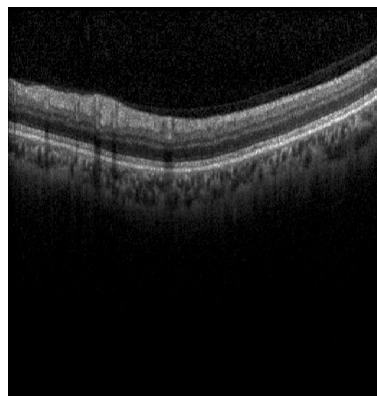


Рис. 3 – Пример В-скана, полученного ОКТ-методом

В дальнейших исследованиях было принято решение опираться на базу изображений проекта STARE и базы ОКТ-изображений, т.к. STARE является размеченной базой, которую можно использовать для машинного обучения, а базы ОКТ-изображений позволяют получать актуальные изображения сетчатки современным неинвазивным методом.

1. STructured Analysis of the Retina [Электронный ресурс] / CLEMSON CECAS. – Юж. Калифорния, 2018. – Режим доступа: <http://cecas.clemson.edu/ahoover/stare/>. – Дата доступа: 16.09.2018.
2. Hoover, A. Locating Blood Vessels in Retinal Images by Piece-wise Threhsold Probing of a Matched Filter Response /A. Hoover, V. Kouznetsova, M. Goldbaum // IEEE Transactions on Medical Imaging – 2000. – Vol. 19, № 3. – P. 203–210.
3. Hoover, A. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels /A. Hoover, M. Goldbaum // IEEE Transactions on Medical Imaging – 2000. – Vol. 22, № 8. – P. 203–210.
4. Fundus Fluorescein Angiogram Photographs of Diabetic Patients [Электронный ресурс] / IUMS. – Эсфахан, 2014. – Режим доступа: <https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1/fundus-fluorescein-angiogram-photographs-of-diabetic-patients>. – Дата доступа: 16.09.2018.
5. Hajeb, SH. A New Combined Method Based on Curvelet Transform and Morphological Operators for Automatic Detection of Foveal Avascular Zone /SH. Hajeb, H. Rabbani, MR. Akhlaghi// Signal, Image and Video Processing (Springer) – 2014. – Vol. 8, № 2. – P. 205–222.
6. OCT data and Color Fundus Images [Электронный ресурс] / IUMS. – Эсфахан, 2014. – Режим доступа: <https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1/oct-fundus-right-left>. – Дата доступа: 16.09.2018.

МОДЕЛИРОВАНИЕ СИСТЕМЫ УПРАВЛЕНИЯ ТОРМОЖЕНИЕМ МОБИЛЬНОЙ РОБОТИЗИРОВАННОЙ ПЛАТФОРМЫ

Татур М. М., Сасин Е. А., Сидорович А. С.

Кафедра электронных вычислительных машин, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: a.s.sidorovich@gmail.com

В работе рассмотрена постановка учебной задачи нелинейного управления и варианты ее решения, которые предназначены для использования в университетских курсах по интеллектуальной робототехнике.

ВВЕДЕНИЕ

При проектировании робототехнических комплексов типовой задачей является задача нелинейного управления исполнительными устройствами. Нелинейность может проявляться как следствие адаптивности принятия решений в тех случаях, когда система в реальном времени реагирует на свое состояние и/или на состояние внешней среды. При изучении (или при проектировании) алгоритмов нелинейного управления возникает определенная проблема. Так, изучаемый алгоритм можно исследовать (или демонстрировать) только в динамике и только в условиях взаимодействия объекта управления с внешней средой, но использовать реальный объект или реальную внешнюю среду, особенно на ранних стадиях проектирования, практически невозможно. Поэтому, применяют компьютерное моделирование с использованием современных универсальных систем, таких как MathLab, MathCAD и др., либо специально разработанных программ. При моделировании важно не только владеть аппаратом моделирования, но и корректно ставить модельные задачи. Именно этому посвящается настоящая работа. А поскольку рассматриваемая задача относится к разряду учебных, то к ней предъявляются дополнительные требования, такие как наглядность и простота: чтобы обучаемые могли интуитивно понимать суть методологии без обращения к сложным математическим выкладкам или специальным знаниям.

I. ПОСТАНОВКА ЗАДАЧИ

Пусть платформа движется с начальной скоростью V_0 , и на расстоянии S_0 от препятствия включается автомат торможения. Платформа должна плавно остановиться в непосредственной близости от препятствия. Алгоритм управления торможением **A** постоянно получает текущие значения скорости v и расстояния S , оставшегося до препятствия, и вычисляет усилие F , прикладываемое к системе торможения (Рис.1).

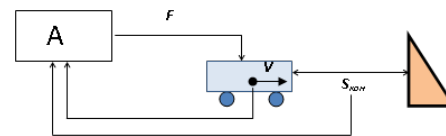


Рис. 1 – Схема упрощенной математической модели процесса торможения мобильной платформы

Требуется предложить алгоритм управления торможением, разработать математическую модель внешней среды и провести серию модельных экспериментов, в ходе которых оценить качество управления.

II. МОДЕЛЬ ВНЕШНЕЙ СРЕДЫ

Решение задачи следует начинать с построения модели внешней среды. На начальном этапе исследования, упростим модель и абстрагируемся от массы платформы, инерционных эффектов, инструментальных ошибок измерений параметров и др. факторов, объективно влияющих на процесс торможения. В идеальном случае будем полагать, что сила F и замедление a — прямо пропорциональны, т.е. $F \sim a$ и мы, теоретически, можем «обеспечить» любое необходимое замедление. При этом, обучаемым предлагается опираться на знания средней школы, т.е. оперировать категориями: скорость, расстояние, время, ускорение (замедление), а для расчета текущих значений V и S в модели «внешней среды» использовать известные формулы замедленного движения. Тогда общую модель можно представить, как показано на рисунке 2.

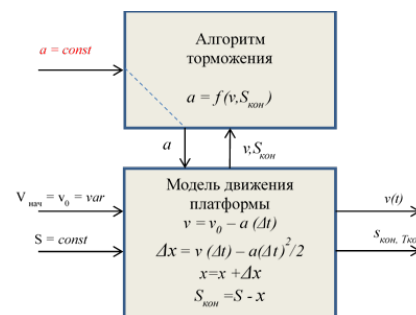


Рис. 2 – Схема упрощенной математической модели процесса торможения мобильной платформы

На данном рисунке в явном виде модель разделена на две части. Первая — «Модель движения платформы», которая рассчитывает оставшееся расстояние до препятствия и текущую скорость, является моделью внешней среды. Вторая часть — представляет собой алгоритм управления торможением.

III. АЛГОРИТМЫ УПРАВЛЕНИЯ ТОРМОЖЕНИЕМ

Качество управления торможением, по сути, является субъективным понятием и может трактоваться достаточно широко [1-3]. Чтобы перейти к конкретным рассуждениям предлагается рассмотреть несколько вариантов (или режимов) торможения и изобразить семейство соответствующих графиков зависимости скорости от времени торможения (Рис.3.).

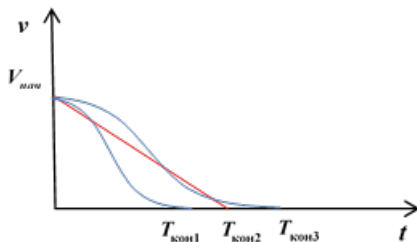


Рис. 3 – Семейство графиков, характеризующих различные варианты торможения

На рисунке красным цветом изображен график равнозамедленного движения. Понятно, что данный вариант не обеспечивает плавности остановки в начальной и конечной точках. Также очевидно, что данный режим не обеспечит адаптивность торможения в зависимости от внешних факторов. Также на рисунке приведены возможные варианты режимов торможения, при которых обеспечивается плавность остановки, причем конечное время может быть как большим T_3 , так и меньшим T_1 , по сравнению равнозамедленным (простейшим) вариантом. Каждому из вариантов будет соответствовать свой алгоритм нелинейного управления. Если выбрать один из вариантов в качестве эталонного, тогда под качеством управления будем понимать расхождение графиков реализуемого и эталонного режимов. Количественно, это расхождение может быть оценено известными способами в результате компьютерного моделирования системы в целом [4]. При необходимости несложно для каждого режима торможения получить график зависимости S от времени.

В простейшем случае, для построения алгоритмов управления могут быть применены эвристические подходы, основанные аппроксимации графика эталонного варианта торможения линейными и/или нелинейными функциями. В дальнейшем, при более глубоком изучении теории управления для разработки алгоритмов могут быть использованы специальные знания из теории регуляторов (ПИД-регуляторы, нечеткие контроллеры) [5].

ЗАКЛЮЧЕНИЕ

В настоящей работе предложена учебная задача моделирования нелинейного процесса управления торможением. Очевидно, что требование $S \rightarrow \min$ напрямую связано с вероятностью аварии, т.е. чем ближе $S \rightarrow 0$, тем вероятнее случай столкновения платформы с препятствием. В работе этот аспект не рассматривался. Также, очевидно, что требование $T \rightarrow \min$ входит в противоречие с требованиями $S \rightarrow \min$ и плавности останова. Всегда можно задать достаточно малое время, за которое обеспечить плавную остановку реальной системы станет невозможным. Оптимизация принятия решений по нескольким критериям также не рассматривалась. При кажущейся простоте настоящая задача отражает методологию моделирования в целом и может быть усложнена за счет специальных алгоритмов управления и детализирована относительно влияющих факторов внешней среды. Также очевидно, что по аналогии можно сформулировать и решать задачи моделирования в других приложениях, например управлением поворота, управлением акселератором и т.п.

1. Hedrick, J. K. Brake System Modeling and Control. / J. K. Hedrick, M. Uchanski//Berkeley, USA/ –2001.
2. Ким, Д. Алгебраические методы синтеза систем автоматического управления / Д. Ким.//–М.: Физматлит. –2014.
3. Brake System Modeling, Control And Integrated Brake/throttle Switching Phase I. University of California. /J. K. Hedrick [and others].//Berkeley, USA/ –1997.
4. Ahmad, H. A. Dynamic Braking Control for Accurate Train Braking Distance Estimation under Different Operating Conditions / H. A. Ahmad, // Blacksburg, VA, USA/ –2013
5. TRAN Development of a Tele-guidance System with Fuzzybased Secondary Controller / M. D. PHUNG [and others].//2010 11th. Int. Conf. Control, Automation, Robotics and Vision Singapore, 5-8th December 2010/ – 2010.

АВТОМАТИЗАЦИЯ КОНТРОЛЯ ЗНАНИЙ КАК МЕТОД ОПТИМИЗАЦИИ ПРОЦЕССА ОБУЧЕНИЯ

Савчук А. А., Нестеренков С. Н.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: savchuk.andrei.dev@gmail.com, nsn@bsuir.by

В этой статье рассматриваются основные виды тестирования знаний, приводятся их главные преимущества, а также производится их сравнение. Представлены возможные проблемы в области контроля знаний. В качестве одного из решений предлагается автоматизировать процесс тестирования знаний. Описаны причины, преимущества и возможные результаты автоматизации процесса тестирования знаний.

ВВЕДЕНИЕ

Контроль знаний является важной составляющей процесса обучения. Целью контроля является определение качества усвоения учащимися программного материала, диагностирование и корректирование их знаний и умений, воспитание ответственности к учебной работе. Систематизированный контроль позволяет определить проблемные места в программе и методике обучения, благодаря чему можно провести процесс оптимизации.

I. Виды контроля

На практике применяется 4 вида контроля знаний:

1. Устный контроль - наиболее распространенный метод контроля знаний студентов. При устном контроле устанавливается непосредственный контакт между преподавателем и студентом, благодаря которому преподаватель имеет возможность изучить индивидуальные особенности усвоения материалов студентом.
2. Письменный контроль - наряду с устной является важнейшим методом контроля знаний, умений и навыков студентов. Однородность заданий ставит всех студентов в равные условия. При письменном контроле за достаточно короткое время можно определить уровень знаний большого количества учащихся.
3. Практический контроль – позволяет выявить, как студенты умеют применять полученные знания на практике, насколько они овладели необходимыми умениями, главными компонентами деятельности. В процессе выполнения профессиональных заданий студент обосновывает принятые решения, что позволяет установить уровень усвоения теоретических положений, т.е. одновременно с проверкой умений осуществляется проверка знаний.
4. Стандартизированный контроль – предназначен для большой выборки учащихся. Обычно, стандартизированный контроль

предполагает тестирование. Тестирование позволяет определить объем, время и уровень усвоения поданной информации, умение использовать и перерабатывать полученную информацию. Благодаря ему за единицу времени можно проверить большую группу людей и затратить меньше времени на проверку по сравнению с письменной формой. Оно имеет некоторые особенности, направленные на увеличение точности измерений, такие как: сложные методы расчета результатов, алгоритмы формирования вопросов и вариантов ответов, системы шкалирования вопросов.

II. ПРОБЛЕМЫ И ПРЕИМУЩЕСТВА

При устном контроле за единицу времени можно проверить лишь одного студента, при письменном же возможно проверять большее количество учащихся, но на расчет результатов может уйти много времени. Помимо этого оно требует расходов на печатные материалы, а также повышает нагрузку на преподавателя [1]. Самым оптимальным и объективным решением выступает тестирование. Однако письменное проведение тестирования становится достаточно накладным процессом, если использовать все ее особенности. Это ставит актуальным вопрос использования вычислительной техники для проведения тестирования или, другими словами, автоматизации этого процесса. Благодаря развитию информационных технологий это стало возможным.

Преимуществами автоматизации тестирования знаний, которые позволяют сократить время преподавателя на проведение тестов, являются ее оперативность и технологичность обработки данных тестирования, автоматизация процедуры педагогического контроля, когда обучаемый выполняет задание в непосредственном диалоге с ЭВМ, результаты сразу переносятся в блок обработки, что позволяет за довольно короткий срок провести процесс дифференциации знаний большого количества испытуемых [2].

Использование вычислительной техники для расчета результатов дает возможность использовать не только дихотомическую оценку, но и политомическую, что увеличивает точность оценки знаний.

Существующие технологии позволяют создавать программное обеспечение, которое без вмешательства пользователя будет оптимизировать расчет результатов тестирования. На вход такой программе будут поступать результаты тестов студентов, которые в процессе анализа, к примеру, будут использоваться для переформирования оценок за задания.

Развитие искусственного интеллекта позволит минимизировать вмешательство пользователя в процесс создания набора тестовых заданий. Благодаря ему можно автоматизировать следующие задачи:

1. Разбиение теста на варианты равной сложности. Достаточно тяжело без использования вычислительной техники разбить набор вопросов на варианты равной сложности. Эту задачу можно легко решить благодаря целочисленному программированию.
2. Переопределение веса задания. В ситуациях, когда на одно задание отвечают все испытуемые, имеет смысл уменьшить вес данного задания.

Централизованное хранение тестов и их результатов позволяет за короткое время получить в удобном для чтения виде информацию, которую можно использовать для составления статистики, анализа результатов, который позволяет и преподавателям и студентам определить уровень подготовки студента с оценкой его результата, эффективно использовать при подготовке к зачету или экзамену варианты тестов, возможности выявить недостаточно изученную тему, контролировать текущую успеваемость в семестре, определять рейтинги студентов по учебной дисциплине.

Помимо оценки студентов, благодаря автоматизации процесса контроля знаний, администрации учреждения образования представляется возможность в удобном виде следить за работой преподавателей, а именно: оценить уровень освоения материалов конкретного преподавателя студентами; получить оценку работы преподавателя, направленной на улучшение результатов (оптимизацию процесса) обучения студентов, за определенный промежуток времени.

Также грамотно построенная система контроля знаний позволит расширить возможности для планирования и распределения учебной нагрузки преподавателей [3].

Все испытуемые находятся в одинаковых условиях, как по содержанию заданий, так и во временных рамках, что позволяет исключить

субъективизм и предвзятость со стороны преподавательского состава и, в тоже время, оценить полученные теоретические знания в полном объеме учебной программы. Этот факт также позволяет студентам объективно видеть результаты своего обучения и избежать ситуаций, когда студент не согласен с оценкой, которую ему предлагает преподаватель, что может снизить стимуляцию к обучению и, как следствие, ухудшить успеваемость.

Основные затраты при автоматизации процесса тестирования знаний приходятся на разработку качественного инструментария, т.е. имеют разовый характер. Затраты на проведение тестирования с использованием программных средств значительно ниже, чем при письменном тестировании.

Таким образом, автоматизация тестирования знаний является универсальным решением следующих проблем в сфере контроля знаний:

1. Затрачиваемое время на проведение тестирования и расчетов результатов;
2. Хранение результатов тестирования и получение статистики за определенный промежуток времени;
3. Субъективность преподавателей;
4. Управление образовательным процессом [4];
5. Финансовые затраты.

1. Нестеренков, С. Н. Алгоритм распределения учебной нагрузки между преподавателями кафедры / С. Н. Нестеренков // Информационные технологии и системы – 2015 (ИТС 2015) : материалы междунар. науч. конф., Минск, 28 октября 2015 г. / Белорус. гос. ун-т информатики и радиоэлектроники; редкол.: Л. Ю. Шилин [и др.]. – Минск, 2015. – С. 272–273.
2. Савчук, А. А. Современные тенденции в организации автоматизированного контроля знаний обучающихся / А. А. Савчук, С. Н. Нестеренков // Дистанционное обучение – образовательная среда XXI века : материалы X Междунар. науч. –метод. конф., Минск, 7–8 декабря 2017 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Б. В. Никульшин [и др.]. – Минск, 2017. – С. 173.
3. Нестеренков, С. Н. Методика планирования и распределения учебной нагрузки преподавателей кафедры в высшем учебном заведении / С. Н. Нестеренков, Б. В. Никульшин // Информационные системы и технологии (ИСТ – 2013) : материалы XIX Междунар. науч. –техн. конф., Нижний Новгород, 19 апреля 2013 г. / Нижегород. гос. техн. ун-т. – Н. Новгород, 2013. – С. 341.
4. Нестеренков, С. Н. Интегрированная информационная система как средство автоматизации управления образовательным процессом в учреждениях высшего образования / С. Н. Нестеренков, Т. А. Рак, О. О. Шатилова // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. – Минск, 2017. – С. 212.

СХЕМА ОФЛАЙН-ГЕНЕРАЦИИ ЭФЕМЕРНЫХ КЛЮЧЕЙ ШИФРОВАНИЯ В ОДНОРАНГОВОЙ СЕТИ

Захарченко К. В., Шилин Л. Ю.

Кафедра информационных технологий автоматизированных систем, кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: cvzakharchenko@gmail.com

В статье предлагается возможный вариант расширения протокола Signal для выполнения в условиях отсутствия центрального сервера с сохранение свойств протокола. В частности предлагается альтернативный метод распространения эфемерных ключей шифрования, основывающийся на использовании криптографических билинейных отображений и сложности решения задачи суммы подмножества.

ВВЕДЕНИЕ

Во многих задачах распределённые пиринговые сети являются выгодной альтернативой клиент-серверным решениям. Причинами этому являются: отсутствие затрат на поддержку и разработку серверной части, автоматическая масштабируемость сети за счёт использования ресурсов клиентов, отсутствие единой точки отказа и прочие. Но помимо решения многих проблем клиент-серверной архитектуры, пиринговые сети также ставят перед разработчиками ряд новых задач. Так, например, при реализации распределённой системы мгновенного обмена сообщениями особой задачей стоит протокол шифрования сообщений. В клиент-серверных решениях одним из наиболее распространённых и заслуживших доверие протоколов является протокол Signal, использующийся в таких системах обмена сообщениями, как WhatsApp, Facebook Messenger, Google Allo и других [1]. Целью данной работы стоит адаптация протокола Signal к работе в условиях отсутствия центрального сервера.

I. КРАТКОЕ ОПИСАНИЕ ПРОТОКОЛА SIGNAL

Протокол Signal использует 10 различных классов ключей и его описание достаточно объёмно, подробный анализ протокола представлен в литературе [1]. Список полезных свойств этого протокола включает в себя: сквозное шифрование, прямая и обратная секретность (forward and backward secrecy), генерация сессионного ключа в режиме офлайн (пользователи могут обмениваться сообщениями не присутствуя в сети одновременно). В начале выполнения протокола генерируются асимметричные ключи 3 типов:

1. долгосрочный ключ;
2. кратковременный ключ;
3. набор эфемерных ключей.

Открытая часть долгосрочного и кратковременного ключа публикуется, закрытая часть хранится пользователем в секрете. Долгосрочный ключ используется для подписи и аутентификации как сообщений, так и кратковременных

и эфемерных ключей. Кратковременный ключ используется для генерации сессионного ключа, для чего протокол Signal утилизует метод двойного храповика (double ratchet). Эфемерные ключи используются для обеспечения дополнительных свойств безопасности и в отличие от первых двух ключей требуют особый метод публикации, реализованный на центральном сервере. Если пользователь A желает отправить сообщение пользователю B , он обращается к серверу с запросом эфемерного ключа, сервер выдаёт ключ из списка ключей пользователя B и удаляет ключ из списка. После того, как эфемерные ключи заканчиваются на сервере, они перестают использоваться в протоколе Signal и протокол теряет свойства прямой и обратной секретности.

II. ПОЛУЧЕНИЕ ЭФЕМЕРНЫХ КЛЮЧЕЙ В РАСПРЕДЕЛЁННОЙ СЕТИ

Протокол Signal требует, чтобы сервер, хранящий эфемерные ключи не выдавал каждый ключ более, чем одному пользователю, то есть удалял уже использованные ключи. Такая функциональность трудно реализуема в условиях одноранговой пиринговой сети без участия арбитров с особыми привилегиями. Сложно доверить пользователям распределённой сети возможность удаления чужих ключей. Поэтому в данной работе предлагается производить генерацию эфемерного ключа стороной, отправляющей сообщение, но так, чтобы только принимающая сторона могла восстановить эфемерный ключ за время полиномиально зависящее от длины ключа.

III. ИСПОЛЬЗУЕМЫЕ КОНСТРУКЦИИ

Для построения целевой схемы в данной работе используются криптографические билинейные отображения и свойства NP -полной задачи суммы подмножества. Дадим краткие определения.

Пусть имеется n групп $\{H_1, \dots, H_n\}$ с генераторами $G = \{g_1, \dots, g_n\}$ соответственно, тогда семейство билинейных отображений на этих груп-

пах задаётся следующим образом[3]:

$$\{e_{i,j} : H_i \times H_j \rightarrow H_{i+j} | i, j \in [n]; i + j \leq n\},$$

где для каждого $e_{i,j}$ выполняется

$$e_{i,j}(g_i^a, g_j^b) = g_{i+j}^{ab} : \forall a, b \in Z_p.$$

Задача суммы подмножества, заданная множеством целых чисел $V = \{v_1, \dots, v_n\} \subset Z_p$ и числом t , состоит в том, чтобы найти такое подмножество $T \subseteq V$, что $t = \sum_{v \in T} v$. Эта задача является NP -полной и на данный момент неизвестен алгоритм эффективно решающий эту задачу в общем случае[2].

IV. СХЕМА ГЕНЕРАЦИИ ЭФЕМЕРНЫХ КЛЮЧЕЙ

Предлагаемая схема описана в данном разделе. Пусть в качестве открытых параметров протокола выступают: параметр k , задающий уровень безопасности системы, параметр $n = (k(k+1))/2$, задающий размер семейства билинейных отображений и семейство билинейных отображений $\{e_{i,j} | i, j \in [n]; i + j \leq n\}$.

Пользователь A , желающий получать сообщения, рандомизированным образом генерирует множество чисел $V = \{v_1, \dots, v_k\} \stackrel{R}{\subset} Z_p$, секретный параметр $\alpha \stackrel{R}{\leftarrow} Z_p$ и генераторы групп $G = \{g_1, \dots, g_n\}$. Далее пользователь кодирует множество V следующим образом:

$$P = \{p_1, \dots, p_k\} = \{g_1^{\alpha^{v_1}}, \dots, g_k^{\alpha^{v_k}}\}.$$

В качестве открытой информации протокола пользователь A публикует параметры V и P , а локально сохраняет генераторы групп G и секретный параметр α .

Пользователь B , желающий отправить сообщение пользователю A , получает из распределённой сети параметры $V = \{v_1, \dots, v_k\}$ и $P = \{p_1, \dots, p_k\}$. Для генерации эфемерного ключа пользователь B выбирает случайное подмножество $I = \{i_1, \dots, i_m\} \stackrel{R}{\subset} [k]$ так, что $|I| \geq 2$ и вычисляет параметры:

$$kv = \sum_{i \in I} v_i,$$

$$ki = \sum_{i \in I} i,$$

$$\begin{aligned} kp &= f_m(p_{i_1}, \dots, p_{i_m}) = f_m(g_{i_1}^{\alpha^{v_{i_1}}}, \dots, g_{i_m}^{\alpha^{v_{i_m}}}) = \\ &= f_m(g_{i_1}, \dots, g_{i_m})^{\alpha^{v_{i_1} + \dots + v_{i_m}}} = g_{ki}^{\alpha^{kv}}, \end{aligned}$$

где $f_m : H_{i_1} \times \dots \times H_{i_m} \rightarrow H_{ki}$ – рекурсивная функция

$$\begin{aligned} f_m(p_{i_1}, \dots, p_{i_m}) &= \\ &= e_{i_1 + \dots + i_{m-1}, i_m}(f_{m-1}(p_{i_1}, \dots, p_{i_{m-1}}), p_{i_m}), \end{aligned}$$

с базой рекурсии

$$f_2(p_{i_1}, p_{i_2}) = e_{i_1, i_2}(p_{i_1}, p_{i_2}).$$

Фактически функция f_m выполняет свёртку слева списка p_{i_1}, \dots, p_{i_m} по билинейным отображениям $e_{i,j}$.

Полученные параметры kv и ki подписываются долгосрочным ключом и пересылаются пользователю A в открытом виде через распределённую сеть, а параметр kp используется при шифровании сообщения, как эфемерный ключ.

Принимающая сторона, получив параметры kv и ki , может восстановить использованный при шифровании эфемерный ключ, так как ей известны генераторы групп и секретный параметр α . Эфемерный ключ можно вычислить как

$$kp = g_{ki}^{\alpha^{kv}}.$$

Таким образом происходит генерация и восстановление эфемерного ключа. Третья сторона C , перехватывающая данные, посылаемые по открытому каналу имеет доступ к параметрам V , P , ki и kv . Чтобы восстановить ключ kp , сторона C должна решить задачу нахождения подмножества $T \subseteq V$, для которого верно $\sum_{v \in T} v = kv$, что эквивалентно решению задачи суммы подмножества. Для решения этой задачи в общем случае (при случайно сгенерированных v_i) количество операций, которое необходимо выполнить лучшему из известных алгоритмов[2], экспоненциально зависит от размера множества V , что на практике при достаточном размере V делает задачу нерешаемой за какое-либо осмысливаемое время. Более того, даже, если третья сторона сможет решить один экземпляр задачи, это не даст ей возможность решать другие экземпляры задачи с каким-либо преимуществом, в чём и заключается эфемерность полученного ключа.

ЗАКЛЮЧЕНИЕ

Статья предлагает возможный вариант использования криптографических билинейных отображений для адаптации выполнения протокола Signal в условиях одноранговой сети. Преимущество предложенной схемы заключается в отсутствии необходимости в арбитраже со стороны центрального сервера. Недостатки отражаются в увеличении сложности протокола и объёма вычислений. Предложенная схема основывается на трудности решения NP -полной задачи. Поэтому злонамеренная сторона может рассекретить сгенерированный эфемерный ключ, но для этого ей понадобится огромное количество вычислительных ресурсов, экспоненциально зависящее от длины ключа. Дальнейшую работу следует направить на изучение точных временных затрат, связанных с выполнением протокола, объёма передаваемых через распределённую сеть данных и на тщательный анализ параметров безопасности предложенной схемы.

1. Cohn-Gordon K. et al. A formal security analysis of the signal messaging protocol //Security and Privacy (EuroS&P), 2017 IEEE European Symposium on. – IEEE, 2017. – С. 451-466.
2. Cormen T. H. et al. Introduction to algorithms. – MIT press, 2009. – С. 1128-1133
3. Boneh D., Franklin M. Identity-based encryption from the Weil pairing //Annual international cryptology conference. – Springer, Berlin, Heidelberg, 2001. – С. 213-229.

ИСПОЛЬЗОВАНИЕ МНОГОУРОВНЕВОЙ МОДЕЛИ ДЛЯ ЭФФЕКТИВНОГО УПРАВЛЕНИЯ ДАМБОЙ И ПРЕДСКАЗАНИЯ НАВОДНЕНИЙ

Шлеменков А. А., Гусак Я. О.

Факультет компьютерных систем и сетей, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {alex.shlemenkov, yaninagusak}@gmail.com

В данной работе будет предложена многоуровневая модель, которая значительно улучшает качество предсказаний потока в дамбе по сравнению со стандартными подходами.

ВВЕДЕНИЕ

Наводнение – одно из наиболее разрушительных стихийных бедствий. Для минимизации вероятности необратимых последствий необходимо не только как можно более точно предсказывать поток в дамбе, но и делать это заранее. Такие предсказания помогут скорректировать управление дамбой и избежать катастрофы. Другой целью моделирования служит оптимизация производства электроэнергии дамбой.

I. ОПИСАНИЕ ДАННЫХ

Данные задачи представляют собой временные ряды со значениями потока через дамбу в текущий момент, количеством осадков и уровнем воды в метрах на нескольких мостах вверх по течению реки. Также для анализа доступны уровни осадков в мм на площади в 900 км². Эта область почти полностью покрывает бассейн реки, на которой располагается дамба. Форма региона визуализирована на рисунке ниже.

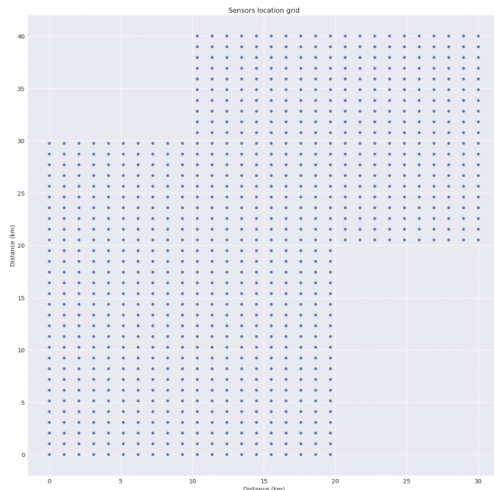


Рис. 1 – Форма области с датчиками

Некоторые виды данных из упомянутых выше не были полными: что-то отсутствовало случайно, у некоторых показателей пропуски были периодическими, поток в дамбе был отрицательным для некоторых промежутков времени.

Убирая из рассмотрения слишком неполные или шумные события, нам удалось получить около 20000 точек и 50 событий. Пример одного из таких событий указан на рисунке.

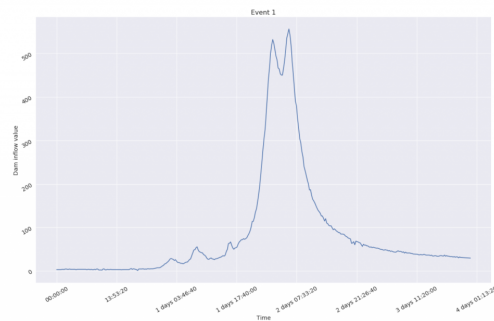


Рис. 2 – Пример экстремального события

II. ПОСТАНОВКА ЗАДАЧИ

Формальная постановка задачи звучит следующим образом: предсказать поток в дамбе через 3 часа. Трехчасовой период времени был выбран из-за того, что этого достаточно для оперативного вмешательства в управления дамбой для предотвращения наводнения. В качестве метрики оценки качества модели предлагается RMSE. В качестве метрики для оптимизации использовалась функция RMSLE [1], которая является хорошим приближением RMSPE, но, в дополнение, еще и хорошо дифференцируется. RMSPE позволяет модели оптимизировать относительную ошибку, а не абсолютную. Т.е, например, для ошибки в предсказании величиной 10 пунктов при реальном значении в 200 и 20 пунктов соответственно, штраф модели не будет одинаковым. Для линейных моделей выбор функционала ошибки является очень важным шагом и сильно влияет на качество предсказаний.

III. ТЕХНИКА ВАЛИДАЦИИ

В качестве техники валидации моделей была выбрана leave-one-out [2] схема, в которой обучение производится на всех кроме одного события, а после оценивается на этом отложенном событии. Дополнительной мотивацией для выбора

такой техники была высокая степень подобию с использованием модели в реальном мире.

IV. МОДЕЛИРОВАНИЕ

1. Для получения базового решения можно использовать линейные модели. В качестве признаков использовались значения всех доступных данных за последние 6 часов. Этого шестичасового окна было достаточно для понимания динамики. Полученные результаты дали среднее значение по схеме leave-one-out в размере 21.12 по метрике RMSE.
2. Другой идеей было использование сверточных сетей для анализа области в 900 км², но, к сожалению, нейросеть не смогла показать достаточную предиктивную силу. После визуального анализа осадков в данной области, не было заметно какой-либо корреляции между потоком и осадками. Низкое качество признаков-осадков подтверждает и тот факт, что коэффициент корреляции между суммой осадков в области и целевой переменной оказался значительно меньше, чем корреляция между текущим значением потока и целевой переменной: 0.6 и 0.825 соответственно.
3. Также была протестирована идея использования RNN, но основным, по нашему мнению, недостатком данных моделей была требовательность к объему данных. Несмотря на то, что GRU намного лучше показывает себя используя меньшее их количество, качество модели оказалась хуже базового, полученного линейной моделью.
4. Использование ансамбля решающих деревьев [3] показало себя хорошо на нашем наборе данных. В терминах величины RMSE результаты были немного лучше базовой модели. Несмотря на это, у методов, которые основаны на решающих деревьях, есть большой недостаток – их невозможность экстраполяции, а в задаче это является очень критичным требованием.

V. ОБУЧЕНИЕ МНОГОУРОВНЕВОЙ МОДЕЛИ

После большого количества экспериментов, было принято решение реализовать комбинированный подход: использовать градиентный бустинг над решающими деревьями на низком значении потока, а линейные модели на высоком значении потока для экстраполяции значения потока. На выбор многоуровневого подхода к финальной модели повлияло и количество данных: малый поток был представлен намного чаще, чем большой. Для того, чтобы использовать трюк с двумя моделями мы обучили SVM [4], который бы разделял точки на два класса: ма-

лый поток и большой поток. Точки из каждой группы подавались в свой регрессор. Конкретное разделяющее значение для малого и большого потока было найдено через множественную процедуру кросс-валидации многоуровневой модели. Для получения конечной модели мы проделали следующие шаги:

1. Разбили данные на N событий.
2. Обучили N SVM классификаторов на N-1 разбиении каждый и предсказали для всех точек на оставшемся событии уровень потока (низкий или высокий).
3. Обучили градиентный бустинг на точках, которым был предсказан низкий уровень потока.
4. Обучили линейную регрессию на точках, которым был предсказан высокий уровень потока.
5. Обучили SVM классификатор на всех N событиях.

VI. РЕЗУЛЬТАТ

Основная идея подхода – это создание многоуровневой модели из двух моделей-специалистов, которые хорошо проявляют себя в своей области. Такая техника позволила снизить ошибку с 21.12 до 10.87 по метрике RMSE. Сравнение с базовой моделью показано на рисунке. На нем видно, что финальная модель реагирует быстрее на выбросы и предсказывает, в целом, точнее.

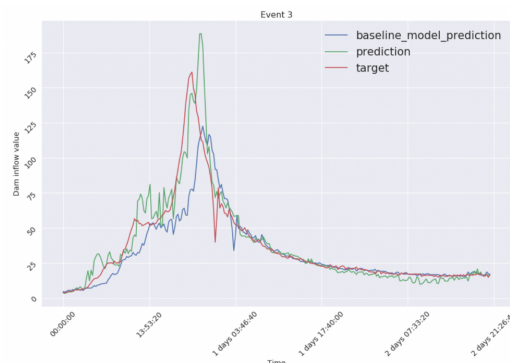


Рис. 3 – Сравнение базовой и многоуровневой моделей

1. RMSLE cost function [Electronic resource]. – <https://www.slideshare.net/KhorSoonHin/rmsle-cost-function> Date of access: 20.09.2018.
2. Evaluation: Leave One Out Cross Validation [Electronic resource]. – Mode of access: <https://www.coursera.org/lecture/predictive-analytics/evaluation-leave-one-out-cross-validation-DfEWO> – Date of access: 20.09.2018.
3. LightGBM [Electronic resource]. – Mode of access: <https://github.com/Microsoft/LightGBM> – Date of access: 20.09.2018.
4. Support Vector Machine – Theory [Electronic resource]. – Mode of access: <https://goo.gl/mTBnV8> – Date of access: 20.09.2018.

МЕТОДИКА СОЗДАНИЯ ВОЗОБНОВЛЯЕМЫХ РАБОЧИХ ПРОЦЕССОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ

Шульга Е. С., Филипук Д. С., Сурков К. А.

Кафедра программного обеспечения информационных технологий, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: egorshulga@outlook.com, filipukd@mail.ru, kirill.surkov@gmail.com

Рассматривается выполнение рабочих процессов в автоматизированных системах. Сбор информации о ходе их выполнения позволяет анализировать успешность выполнения их этапов. После преднамеренного или непреднамеренного отказа выполнение рабочего процесса можно продолжить с этапа, следующего за последним успешно завершённым.

ВВЕДЕНИЕ

Типичным предметом автоматизированной системы масштаба предприятия является рабочий (производственный) процесс. Существуют рабочие процессы, выполняемые целиком на одном компьютере, и существуют рабочие процессы, выполнение которых распределено среди нескольких взаимодействующих компьютеров вычислительной сети. В случае применения распределенных рабочих процессов неизбежно возникают риски по отказам взаимодействующих программных средств и каналов взаимодействия. В связи с этим появляется задача создания возобновляемых рабочих процессов, выполнение которых может быть остановлено на любом этапе и в любой точке, а затем продолжено до успешного завершения.

I. ПОСТАНОВКА ЗАДАЧИ

Далее будем использовать следующие определения:

Рабочий процесс – набор действий для достижения определенной цели. Процессом может быть регистрация пользователя в системе, выполнение определенных вычислений и т.д.

Этап – независимая составная часть рабочего процесса. Этапы объединяются в некоторую последовательность, которая может быть представлена однонаправленным графом без петель. При этом в графе может быть один или несколько начальных, а также один или несколько конечных этапов.

Действие – атомарный элемент процесса, составная часть этапа. Действия могут выполняться последовательно, параллельно, комбинированно или не выполняться вовсе.

Выполнение процессов пользователем можно представить в виде дискретно-детерминированного конечного автомата [1].

Пример процесса приведен на рисунке 1. В данном процессе существует один начальный этап (этап 1), в него входит набор действий: А, В, С, D, Е. Для остальных этапов наборы действий опущены. Примеры ветвления: ветвь этапов 2.1,

3 и этапа 2.2, ветви этапов 5.1, 5.3 и с этапа 5.2 по этап 6. Наличие ветвлений означает, что некоторый компонент автоматизированной системы будет принимать решение, по какому пути графа будет выполняться рабочий процесс. Кроме того, в приведенном примере существует несколько завершающих этапов: этапы 5.1, 5.3 и 6. На данном примере рассмотрим методику создания возобновляемого процесса.

Предполагается, что все компоненты системы имеют возможность хранения некоторой информации. Серверные компоненты могут сохранять ее в базах данных. Клиентская часть сохраняет лишь небольшую часть информации, необходимую для ее работы.

Для внедрения возможности возобновления этапы процесса должны быть независимы друг от друга: обработка действий на каждом этапе должна завершаться сохранением необходимых данных и передачей управления следующему этапу. Данные между этапами могут передаваться через постоянные хранилища (базы данных) или сохраняться в клиентских компонентах автоматизированных систем (при поддержке ими такой возможности).

II. СБОР ИНФОРМАЦИИ О РАБОЧЕМ ПРОЦЕССЕ

По ходу выполнения рабочего процесса необходимо вести сбор данных для анализа времени, успешности выполнения рабочих процессов, этапов и действий, анализ прочей статистики. Кроме того, собранная информация может быть использована для возобновления прерванных рабочих процессов.

Для реализации сбора данных в автоматизированной системе предлагается создание специального компонента и использование постоянного хранилища. Пользователи и другие компоненты могут заниматься выполнением различных рабочих процессов, каждый из которых может содержать различное число этапов с многочисленными действиями. Создание записей о выполнении частей рабочего процесса будет проис-

ходить очень часто, в то время как их чтение, создание отчетов, создание заявок на возобновление рабочих процессов будет происходить намного реже. В связи с этим программный интерфейс компонента сбора данных должен быть максимально упрощён, а именно: он должен содержать единственный метод, который не осуществляет сложных вычислений, а лишь записывает в базу данных переданную ему информацию о выполненном действии.

Все рабочие процессы в автоматизированной системе должны быть описаны в терминах этапов и действий. Эта информация должна быть известна в каждом компоненте. Таким образом, сбор данных о выполненных действиях является сквозной функциональностью (cross-cutting concern) [2].

Для реализации возможности возобновления в каждом рабочем процессе введём дополнительное псевдодействие: «Окончание этапа N». Информация об этом действии должна быть сохранена на каждом этапе как маркер завершения.

Для поддержки ветвления также необходимо сохранять маркерные действия. Если после определенного этапа выполнение рабочего процесса может пойти по одной из веток, то автоматизированная система должна вынести решение об используемой ветке и затем записать через компонент сбора информации о процессе действие «Начат этап M». В примере рабочего процесса на рисунке 1 после выполнения этапа 1 должно быть записано одно из действий: «Начат этап 2.1» или «Начат этап 2.2». Аналогично после этапа 4 сразу же должно быть записано одно из действий: «Начат этап 5.1», «Начат этап 5.2» или «Начат этап 5.3».

III. ВОЗОБНОВЛЕНИЕ РАБОЧЕГО ПРОЦЕССА

Если рабочий процесс прерван, то он может быть возобновлён. Прерывание рабочего процесса может произойти по ряду причин:

- непреднамеренно при отказе одного из компонентов автоматизированной системы;
- непреднамеренно при сбое связи между компонентами;
- преднамеренно при отказе пользователя продолжать рабочий процесс.

Для возобновления прерванного рабочего процесса предлагается следующий алгоритм:

1. Найти последний этап с записанным маркерным действием «Окончание этапа S_i ».
2. Если после этапа S_i следует единственный этап S_{i+1} , то он и является следующим для выполнения. Процесс нужно продолжить именно с него.
3. Если после этапа S_i следует несколько этапов S_{i+1}^k , то необходимо проверить, есть ли записанные действия в каждом из них.
4. Искомым этапом для возобновления процесса является первый с непустым множеством записанных этапов.
5. Если ни для одного из этапов S_{i+1}^k не записано действий, то необходимо вернуться на один этап назад. Этапом для восстановления процесса будет S_i .

Запуск механизма возобновления может происходить различными способами. В случае отказов серверные компоненты могут периодически совершать опрос на предмет восстановления функционирования других компонентов автоматизированной системы. В случае преднамеренной остановки процесса пользователем можно отправить ему уведомление с предложением закончить выполнение рабочего процесса.

IV. ЗАКЛЮЧЕНИЕ

Описанная методика позволяет создавать возобновляемые рабочие процессы в автоматизированных системах. Возобновление выполняется после преднамеренных или непреднамеренных отказов. Предлагается внедрение специального компонента по сбору информации о выполнении процессов в системе, который способен обработать большое число одновременных заявок на сохранение данных. В случае прерывания рабочего процесса система не оказывается в неконсистентном состоянии, и выполнение рабочего процесса можно продолжить на следующем этапе за последним успешно завершённым этапом.

V. СПИСОК ЛИТЕРАТУРЫ

1. Моделирование систем / Б. Я. Советов, С. А. Яковлев. – М.: Высш. шк., 2001. – 343 с.
2. Kiczales, G. Aspect-Oriented Programming / G. Kiczales, J. Lamping // Proceedings of the 11th European Conference on Object-Oriented Programming (ECOOP 1997) – Jyväskylä, Finland, 1997. – P. 220–242.

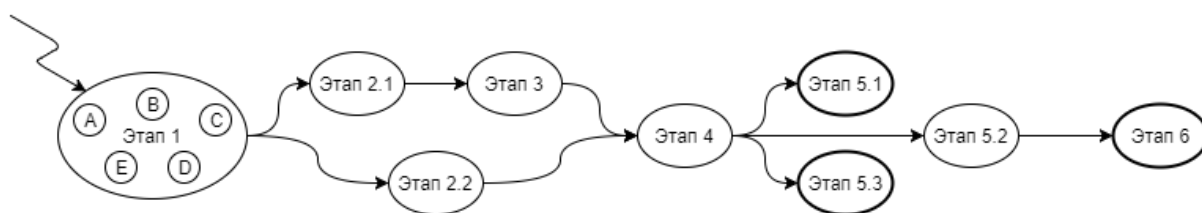


Рис. 1 – Пример процесса

ПРОГНОЗНЫЕ ЗАДАЧИ В CRM-СИСТЕМАХ

Шуляк Е. В., Гуринович А. Б., Смирнов В. Л.

Кафедра ИТАС, Кафедра ВМиП, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: yegorshulak@gmail.com, gurinovich@bsuir.by

Оценка качества обслуживания клиентов, анализ предпочтений и прогнозирование представляет собой необходимую деятельность по созданию полноценно функционирующей и имеющей высокие показатели системы продаж. В данной работе была использована платформа Cloudera Hadoop и были определены оптимальные модели анализа данных.

ВВЕДЕНИЕ

Для исследования выбраны банковские маркетинговые данные, используемые для решения прогнозных задач в телефонном маркетинге. Они подходят для машинного обучения путем создания единого классификатора. Фактически этот набор данных содержит 150 столбцов, но их количество сокращается до 22 в результате экспертного анализа.

I. ОСНОВНЫЕ ВХОДНЫЕ ПАРАМЕТРЫ

Информация представлена в CSV – текстовом формате для представления табличных данных. Основные входные параметры следующие.

- Данные банковских клиентов:
 - age_job: стаж работы;
 - categorical_job: тип работы;
 - marital: семейное положение;
 - education: образование;
 - default: наличие кредита в принципе;
 - housing: наличие кредита на жильё;
 - personal_loan: наличие личного кредита.
- Данные, связанные с последними контактами текущей кампании:
 - contact: тип связи (сотовая или стационарная);
 - month: последний месяц звонка;
 - day_of_week: последний день недели звонка;
 - duration: Продолжительность последнего звонка в секундах.
- Другие атрибуты:
 - campaign: количество контактов участников на данный момент;
 - rdays: количество дней, прошедших после последнего контакта с клиентом;
 - previous: количество контактов этого клиента, выполненных до данной кампании;
 - routcome: результат предыдущей маркетинговой кампании (по категориям).
- Атрибуты социального и экономического контекста:
 - emp.var.rate: коэффициент изменения занятости – квартальный показатель;

- cons.price.idx: индекс потребительских цен – ежемесячный индикатор;
- cons.conf.idx: индекс доверия потребителей – ежемесячный индикатор;
- euribor3m: Европейская межбанковская ставка предложения – ежедневный индикатор;
- nr.employed: индекс занятости – квартальный показатель;
- Выходная переменная (желаемая цель):
 - y – Подписался ли человек на срочный депозит? ('да', 'нет')

Необходимо объяснить, что этот набор данных предоставлен исследованием Моро и др.[1] об успехах или неудачах при рекламе долгосрочных кредитов в банках. Он состоит из данных, собранных с 2008 по 2013 год. В нём был использован полуавтоматизированный выбор функций для изучения лучшего моделирования. «Полуавтоматизированным» его можно назвать потому, что на некоторых шагах были использованы экспертные суждения, основанные на интуитивном знании бизнеса, в частности, было использовано 14 вопросов для группировки важности параметров.

II. АНАЛИЗ

Итак, после предварительной обработки данных, набор чист. Мной будут использованы 4 модели интеллектуального анализа данных:

Логическая регрессия (LR) – статистическая модель, используемая для предсказания вероятности возникновения некоторого события путём подгонки данных к логистической кривой.[2]

Дерево принятия решений (DT) – средство поддержки принятия решений, использующееся при анализе данных для прогнозных моделей. Структура дерева представляет собой «листья» и «ветки». На «ветках» записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах – атрибуты, по которым различаются случаи.

Нейронная сеть (NN) – математическая модель, а также её программное воплощение, построенная по принципу организации и функцио-

нирования биологических нейронных сетей – сетей нервных клеток живого организма.[3]

Метод опорных векторов (SVM) – набор схожих алгоритмов использующихся для задач классификации и регрессионного анализа. Идея метода – перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей классы. Разделяющая гиперплоскость максимизирует расстояние до двух параллельных гиперплоскостей. Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора.[4]

Для определения лучшей модели с лучшими параметрами применим два следующих подхода: ROC-анализ (англ. receiver operating characteristic, рабочая характеристика приёмника), использующий график, позволяющий оценить качество бинарной классификации, отображающий соотношение между долей объектов от общего количества носителей признака, верно классифицированных как несущих признак, и долей объектов от общего количества объектов, не несущих признака, ошибочно классифицированных как несущих признак. Количественную интерпретацию метода даёт показатель AUC (англ. area under ROC curve, площадь под ROC-кривой) – площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций.[5] И LIFT-анализ (англ. lift, подъёмная сила), показывающий, насколько повышается вероятность нахождения Т в анализируемом случае, если в нем уже имеется А. Количественную интерпретацию данного метода даёт показатель ALIFT – площадь, ограниченная LIFT-кривой и осью доли ложных положительных классификаций.[6]

III. ГРАФИКИ

На графиках показаны кривые относительно показателей Ложных положительных оценок (FPR) и Правдивых (TPR). Кривая логистической регрессии обладает лучшими параметрами. (см. рис. 1-2)

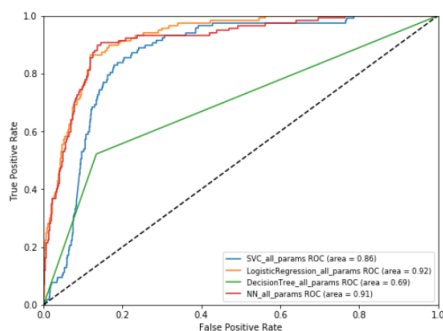


Рис. 1 – ROC-анализ

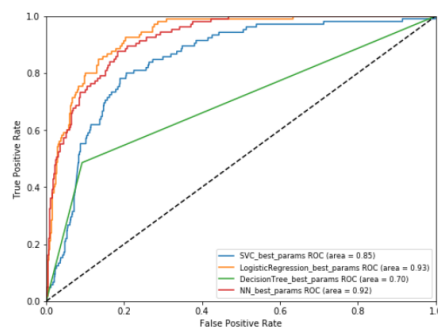


Рис. 2 – LIFT-анализ

«Продолжительность», «возраст», «Европейская межбанковская ставка предложения», «работа» и «образование» занимают первые четыре места в топе показателей важности. Одна из этих переменных – «Европейская межбанковская ставка предложения» за последние три месяца, банком не контролируется. «Продолжительность» сильно влияет на выходную цель (например, если продолжительность = 0, то ответ на вопрос, подписался ли человек на срочный депозит, очевидно, $y = \text{'нет'}$) но, тем не менее, продолжительность не известна до конца вызова, а кроме того, после окончания вызова ответ на вопрос уже известен. Таким образом, если намерение состоит в том, чтобы создать реалистичную прогнозирующую модель, этот параметр лучше вообще не учитывать. Следовательно, основными факторами являются социально-экономические.

IV. ВЫВОДЫ

В исследовании анализировались результаты работы нескольких моделей. До этого проводилась предварительная обработка и выбор функций. Расчёты показали, что оптимальными для решения прогнозных задач в CRM-системах являются логическая регрессия и нейронная сеть.

V. СПИСОК ЛИТЕРАТУРЫ

1. S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. *Decision Support Systems*, Elsevier, 62:22-31, June 2014.
2. Harrell, Frank E. (2001). *Regression Modeling Strategies* (2nd ed.). Springer-Verlag.
3. Паклин Н.Б., Орешков В.И. Глава 9. // Бизнес-аналитика: от данных к знаниям(+CD): Учебное пособие. 2-е изд.. — СПб: Питер, 2013. — С. 428-472.
4. Владимир Вьюгин. Математические основы теории машинного обучения и прогнозирования. — МЦМ-НО, 2013. — 390 с.
5. «A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems». *Machine Learning* 45: 171-186.
6. Tufféry, Stéphane (2011); *Data Mining and Statistics for Decision Making*, Chichester, GB: John Wiley & Sons.

АВТОМАТИЗАЦИЯ ВЕРИФИКАЦИИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СВЯЗАННЫХ С БЕЗОПАСНОСТЬЮ СИСТЕМ

Сивко Б. В.

Лаборатория «Безопасность и электромагнитная совместимость технических средств»,
Белорусский государственный университет транспорта
Гомель, Республика Беларусь
E-mail: bsivko@gmail.com

Рассматривается практика верификации программного обеспечения систем, связанных с безопасностью. В докладе излагается несколько задач верификации, решение которых проводится автоматизированными программными средствами. В качестве таких задач рассматривается определение временных параметров работы устройств, оценка степени диверситета аппаратно-программных комплексов, автоматизация выбора адресов по методу обнаружения отказов на основе доступности адресных данных. Решения основываются на формальной верификации исходного кода программ.

ВВЕДЕНИЕ

Разработка систем, связанных с безопасностью (*safety-critical systems*, ССБ), сопряжена с дополнительными мероприятиями и затратами на обеспечение предъявляемых к ним требований безопасности. Данная задача усложняется для микропроцессорных аппаратно-программных комплексов, так как они обладают высокой сложностью, присущей программному обеспечению (ПО). В то же время, к ССБ предъявляются повышенные требования по безопасности и надежности функционирования, что создаёт потребность в создании эффективных методов и средств для повышения качества решения ключевых проблем разработки и верификации [1].

Одним из способов решения описанной задачи является создание автоматизированных программных средств, позволяющих снизить влияние человеческого фактора и предоставить дополнительный способ полного охвата заданных спецификациями положений. Помимо этого, средства автоматизации могут уменьшить затраты посредством выявления ошибок проектирования на более ранних стадиях разработки — до имитационных испытаний. Для ССБ автоматизация прежде всего рассматривается как дополнительный способ повышения показателей отказоустойчивости и безопасности. Трудности данного направления обусловлены разнообразием решаемых задач и особенностями элементной базы, что осложняет необходимую для автоматизации формализацию.

I. ОЦЕНКА ВРЕМЕННЫХ ПАРАМЕТРОВ

Ряд ССБ относятся к системам реального времени, и поэтому для них характерно требование выполнения определённых временных параметров, например, гарантии перехода в безопасное состояние по заданному тайм-ауту, предоставление возможности задания периода обновления устройств индикации, определение частоты опроса внешних устройств и других [1].

Практика верификации ПО ССБ железнодорожной автоматики и телемеханики показывает, что к основным задачам такого типа относятся: вычисление времени выполнения между двумя произвольными точками, определение обстоятельств заикливания программы, а также доказательство обязательного завершения алгоритма. Кроме того, на практике для циклических систем оказывается актуальным поиск точек программы, где выполнение обязательно произойдёт при каждом выполнении тела цикла. Для автоматического решения данных задач разработано ПО *Formal Time Verifier* [2], с помощью которого возможна верификация программ PIC-контроллеров модели 16F877A, а так как используется общая база команд исполнения PIC, то ПО может применяться для микроконтроллеров других модификаций. Функционально *Formal Time Verifier* проводит синтаксический разбор исходного кода и создаётся граф переходов, и разработанные алгоритмические решения представляют собой решения задач на графах.

II. ОЦЕНКА СТЕПЕНИ ДИВЕРСИТЕТА

Диверситет является одним из основных способов повышения отказоустойчивости и безопасности ССБ, заключающийся в создании как можно более разных систем таким образом, чтобы в случае отказа они повели себя по-разному, что позволяет обнаружить отказ, провести диагностику, перейти безопасное состояние или в режим самовосстановления. Здесь одной из важных задач становится оценка степени полученного диверситета (то есть, различия), позволяющая определить эффективность применяемых методов и средств [1, 3].

Автоматизация решений проблем, связанных с диверситетом, является сложной и актуальной задачей, так как оценка степени достигнутого диверситета на современном этапе производится неформальными методами. Одним из

методов формализации является диверсификация аксиоматических базисов, когда разрабатываемая система опирается на заранее определённые формализованные утверждения [3]. Соответственно, возможна разработка средств автоматизации, позволяющих проверить выполнение данных утверждений для разработанной системы и их истинность во время проявления тех или иных отказов.

Для этих целей разработано ПО *Diverse Axiomatic Basis Checker* [4], позволяющее определить базисы на основе исходного кода программ и информации об микропроцессорной архитектуре. ПО позволяет верифицировать программы для PIC-контроллеров модели 16F877A. Для других моделей или систем необходима адаптация (задание конфигурации) определения базисов. *Diverse Axiomatic Basis Checker* проверяет микропроцессорную систему на константные отказы (*SA*, *stuck-at faults*) и отказы короткого замыкания произвольных информационных линий (*B*, *bridging faults*), в которые входят отказы ячеек памяти, дешифратора команд и выполнения инструкций микроконтроллера, отказы регистров и аккумулятора. Практика применения *Diverse Axiomatic Basis Checker* показала, что подход дополнительно позволяет проверить ряд утверждений, которые вручную проверить затруднительно – например, влияние отказов при рассмотрении больших объёмов памяти. В то же время, ограничениями подхода стали сложность реализации проверки утверждений, в частности, когда для анализа необходима дополнительная информация о работе алгоритма [4, 5].

III. ВЫБОР АДРЕСОВ

Во время применения метода обнаружения отказов на основе доступности адресных данных происходит выбор определённого набора адресов, который зависит от того множества отказов, наличие которых требуется проверить. Идея метода состоит в том, что в случае отказа один из адресов становится недоступным, и на этом основании система может обнаружить проблему, перейти в безопасное состояние или запустить процедуры самовосстановления [6].

У произвольной системы некоторые области адресного пространства заняты исполняемым кодом или данными, и эти области формализуются в виде диапазонов, которые могут быть определены исходя из исходного кода. Во время практического применения требуется, чтобы пользователь мог задать важные для него критерии (множество проверяемых отказов, разрешенные адресные диапазоны) и найти оптимальные варианты. Для решения описанной задачи разработано ПО *Address Detection* [7, 8], реализующее цели автоматизации: уменьшение ошибок во время поиска адресов, нахождение оптимального из возможных наборов и уменьшение затрат. ПО выполняет расчёт в рамках базовых мо-

делей *SA*- и *B*-отказов. Решение для *SA*-отказов является простой задачей, так как они задают конкретные биты адресного регистра. Для одного *B*-отказа задача эквивалентна раскраске графа в два цвета, которая решается за линейное время. Для большего числа *B*-отказов проблема становится *NP*-сложной. *Address Detection* предоставляет решение исходя из практических потребностей, позволяя вычислять пары адресов с минимальным расстоянием друг от друга и с наибольшим минимальным адресом с начала адресного пространства.

ЗАКЛЮЧЕНИЕ

В докладе рассматриваются задачи автоматизации, разработанные алгоритмы и особенности применения предложенных программных средств. ПО *Formal Time Verifier*, *Diverse Axiomatic Basis Checker* и *Address Detection* опробованы в лаборатории «БЭМС ТС» БелГУ-Та и зарегистрированы в 2017 году в Национальном центре интеллектуальной собственности, г. Минск [2, 4, 7].

1. Бочков, К. А. Микропроцессорные системы автоматизации на железнодорожном транспорте : учеб. пособие / К. А. Бочков, А. Н. Коврига, С. Н. Харлап; М-во образования Респ. Беларусь, Белорусский государственный университет транспорта. – Гомель. – 2013.
2. Сивко, Б. В. Formal Time Verifier: свидетельство о регистрации компьютерной программы в Национальном центре интеллектуальной собственности Республики Беларусь № 995 / Б. В. Сивко. – Оpubл. 13.12.2017.
3. Бочков К. А., Разработка отказоустойчивых систем на основе диверсифицированных аксиоматических базисов / К. А. Бочков, С. Н. Харлап, Б. В. Сивко // Автоматика на транспорте: ПГУПС. – 2016. – № 1, т. 2. – С. 47–64.
4. Сивко, Б. В. Diverse Axiomatic Basis Checker: свидетельство о регистрации компьютерной программы в Национальном центре интеллектуальной собственности Республики Беларусь № 996 / Б. В. Сивко. – Оpubл. 13.12.2017.
5. Сивко, Б. В. Автоматизация процесса оценки степени диверсифицированных аппаратно-программных комплексов / Б. В. Сивко // Проблемы безопасности на транспорте : материалы VIII Междунар. науч.-практ. конф., посвящ. Году науки: в 2 ч. Ч. 1, Гомель, 23–24 ноября 2017 г. / БелГУТ; под общ. ред. Ю. И. Кулаженко [и др.]. – Гомель, 2017. – С. 199–200.
6. Сивко, Б. В. Обнаружение отказов на основе доступности адресных данных / Сивко Б.В. // Информационные технологии и системы: материалы междунар. науч. конф., Минск, 26 октября 2016 г. / БГУИР; гл. ред. Л. Ю. Шилин [и др.]. – Минск, 2016. – С. 66–67.
7. Сивко, Б. В. Address Detection: свидетельство о регистрации компьютерной программы в Национальном центре интеллектуальной собственности Республики Беларусь № 983 / Б. В. Сивко. – Оpubл. 23.10.2017.
8. Бочков, К. А. Автоматизация метода обнаружения отказов на основе доступности адресных данных / К. А. Бочков, С. Н. Харлап, Б. В. Сивко // Проблемы безопасности на транспорте : материалы VIII Междунар. науч.-практ. конф., посвящ. Году науки: в 2 ч. Ч. 1, Гомель, 23–24 ноября 2017 г. / БелГУТ; под общ. ред. Ю. И. Кулаженко [и др.]. – Гомель, 2017. – С. 183–184.

ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБРАЗОВАНИИ

Стригалёв Л. С., Тиханович Т. В., Кузнецова О. В.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {orion, tihanovich, ovkuznetsova}@bsuir.by

Рассматриваются возможности интеллектуальных технологий в деле совершенствования образовательной среды методом потенциалов.

ВВЕДЕНИЕ

Сложность современных технологических процессов, возрастающая роль и удельный вес информационных технологий, широко используемых средства искусственного интеллекта, требуют изменений в сфере образования. Сложился разрыв между потребностями практики и возможностями образовательной среды, вызванный запаздыванием внедрения технологий искусственного интеллекта в образование, который обусловлен как сложностью образовательного процесса, так недостаточностью финансирования.

I. ТЕХНОЛОГИИ С ЭЛЕМЕНТАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

По-новому подойти к проблеме индивидуализации обучения позволяет внедрение современных интеллектуальных информационных технологий, в частности, технологий с элементами искусственного интеллекта.

Осуществление идеи личностно-ориентированного обучения, позволяет максимально индивидуализировать учебно-воспитательный процесс, повышает качество обучения и способствует развитию индивидуальных способностей студентов. На это нацелено использование экспертных систем, являющихся одним из элементов искусственного интеллекта.

Одним из перспективных направлений внедрения интеллектуально-информационных технологий в образование является использование прикладных систем на основе методов искусственного интеллекта, таких как экспертные системы, интеллектуальные обучающие системы, экспертные обучающие системы. Интеллектуализация обучающих систем, рост числа программных комплексов, реализующих идеи и принципы искусственного интеллекта, обусловили необходимость в научно-обоснованной методической системе обучения основам искусственного интеллекта будущих педагогов профессионального обучения.

Для преодоления противоречий между постоянно нарастающим объемом знаний, подлежащих усвоению, и реальными сроками подготовки специалистов, возникает необходи-

мость определить новые организационные формы обучения, способы получения знаний обучаемыми, закрепления умений и навыков. Одной из таких форм подготовки и переподготовки высококвалифицированных специалистов в вузах следует считать новую технологию обучения, базирующуюся на применении предметно-ориентированных и проблемно-ориентированных компьютерных обучающих систем.

Использование названных систем способствует повышению эффективности технологий обучения, интенсификации учебного процесса, повышению качества подготовки специалистов.

Необходимость же внедрения искусственного интеллекта в образовательную среду очевидна, поскольку он способен изменить технологии, как основного, так и обеспечивающего контура обучающей системы. Искусственный интеллект «несет индустриализацию» в одну из самых сложных областей человеческой деятельности, повышая тем самым эффективность обучения, снижая затраты на обучение и сроки обучения. Так что знания становятся товаром, к которому применимы известные правила логистики: нужный товар, в нужном месте, нужного количества, нужного качества, в нужное время, с минимальными затратами.

Облачные технологии освобождают студента от технологических, структурных и ресурсных забот, предоставляя ему необходимое технологическое пространство, параметры которого зависят от вида сервиса. Основными видами облачных сервисов (услуг) являются [1]: инфраструктура как сервис (Infrastructure as a Service; IaaS); платформа как сервис (Platform as a Service; PaaS); программное обеспечение как сервис (Software as a Service; SaaS).

Наиболее широкие возможности обеспечивает IaaS, так как в этом случае сервис предоставляется на уровне инфраструктуры. На базе инфраструктуры IaaS реализуются и названные выше и другие сервисы более высокого уровня. Сервисы PaaS обеспечивают доступ к технологическим платформам (к операционным системам и прикладным программам). Самым распространенным видом облачных сервисов является SaaS; по существу, все приложения, которые установ-

лены не у пользователя, а в «облаке» относятся к модели SaaS. Достоинство облачных технологий состоит том, что они позволяют реализовать любые необходимые сервисы и в совокупности с технологиями вебинара создавать эффект полного присутствия с полномасштабным тренингом.

Облачные технологий способны доставить в любую географическую точку требуемые структуру и базовые технологии; сами же эти технологии, находясь в постоянном развитии, под воздействием потребностей практики, постоянно совершенствуются. Так развитие аппаратно-программных средств компьютерных технологий привело к возникновению огромных объемов информации, так называемых больших данных (Big Data), и интернета вещей (Internet of Things, [2]). Возрастающие объемы информации создает не только человек, технологические и научно-исследовательские средства и системы, но и многочисленные мобильные устройства, количество которых значительно превысило население Земли.

Big Data принято характеризовать «четырьмя V»: объем (Volume), разнообразие (Variety), изменчивость (Variability) и скорость (Velocity). Термин Big Data (в русскоязычной среде) обычно используется как для обозначения больших данных (неструктурированных и структурированных), так технологий их обработки, к которым, в частности, относятся технологии Data Mining и машинное обучение (Machine learning), включая глубокое обучение (Deep learning), применяемое для обучения искусственных нейронных сетей. В последние годы термин Big Data потерял актуальность; адекватной заменой ему, как считают многие специалисты, является наука о данных (Data science).

Технологии Data Mining по типу обрабатываемых данных подразделяются: Text Mining (технологии поиска и семантического анализа текста); Web Mining (интеллектуальный анализ данных в Internet); Call Mining («добыча звонков», технология распознавания речи и ее анализ); Audio Mining (извлечение данных из аудиозаписей); Video Mining (извлечение данных из видеозаписей). Технологии Data Mining (развивалась на базе прикладной статистики, теории искусственного интеллекта и машинного обучения) находят широкое применение в многочисленных областях человеческой деятельности.

Искусственные нейронные сети обладают большими комбинаторными возможностями, способны обучаться и самообучаться; являясь узко специализированными, они широко используются в различных сферах человеческой деятельности, в ряде случаев значительно превосходя человека [3].

Внедрение в образование технологий Big Data, Data Mining и нейронных сетей, позволит перестроить образовательные структуры, обострить конкуренцию и поставить слабые вузы в

реальные условия. Возможность же адаптивно обучаться у лидирующих преподавателей лучших вузов по индивидуальным траекториям не только сократит сроки и повысит качество обучения, но и снизит, с использованием технологий прокторинга, издержки, связанные с организацией, и контролем процесса обучения. Современные технологии прокторинга обеспечивают автоматическую идентификацию лица и компьютерного почерка студента, а также позволяют выявлять типичные нарушения при тестировании, присутствие посторонних лиц и голосов.

Наконец, использование интеллектуальных средств, позволит востребованным преподавателям обслуживать не только более широкий контингент обучаемых за счет сокращения затрат на рутинные операции, но и более эффективно совершенствовать свой научный уровень и педагогическое мастерство.

II. Выводы

Технологии искусственного интеллекта применимы как в основном, так и обеспечивающем контуре системы обучения. Главным и ведущим является основной контур, в котором осуществляется формирование компетенций специалистов заданного профиля. Индустриализация этого контура приведет к перестройке и технологий обеспечивающей подсистемы. Некоторые операции окажутся не востребованными или будут модифицированы. Например, часть функций деканата, в частности контроль посещения занятий, успеваемости и других функций могут быть реализованы с использованием технологий прокторинга и Data Mining.

Однако основную цель деятельности деканата по воспитательной и управленческой работе пока современные технологии не смогут заменить, т.к. воспитательная цель заключается в создании условий для формирования профессиональной и социальной компетентности, активной гражданской позиции студентов, способствовать воспитанию личностных качеств будущих специалистов в определённой профессиональной сфере, готовых в новых социально-экономических условиях вносить свой вклад в развитие и укрепление Республики Беларусь.

СПИСОК ЛИТЕРАТУРЫ

1. Проволоцкий, Е. В. Облачные технологии в учебном процессе / В. Е. Проволоцкий, Е. М. Демидок, Л. С. Стригалев, Н. В. Протченко / Информационные технологии и системы 2013 (ИТС 2013): Междунар.научн.-метод.конференции, 23 октября 2013 г. – Минск: БГУИР, 2013. – С. 310–311.
2. Шилин, Л. Ю. Технологии семантической обработки информации в учебном процессе / Л. Ю. Шилин, А. А. Навроцкий, Л. С. Стригалев / BIG DATA and ADVANCED ANALYTICS. – Минск: БГУИР, 2017. – С. 181–183.
3. Шилин, Л. Ю. Нейронные сети в системах обработки данных / Л. Ю. Шилин, А. А. Навроцкий, Л. С. Стригалев / BIG DATA and ADVANCED ANALYTICS. – Минск: БГУИР, 2018. – С. 389–391.

МЕТОДЫ И ИНСТРУМЕНТЫ ОЦЕНКИ ТОЧНОСТИ КОМПЬЮТЕРНЫХ ГЕОЛОГИЧЕСКИХ МОДЕЛЕЙ

Таранчук В. Б., Таранчук В. В

Кафедра компьютерных технологий и систем, Факультет прикладной математики и информатики,
Белорусский государственный университет,

НИИ прикладных проблем математики и информатики Белорусского государственного университета
Минск, Республика Беларусь

E-mail: taranchuk@bsu.by, victoria_t@tut.by

Обсуждаются вопросы использования интегрированного программного комплекса составителя цифровых геологических, геоэкологических моделей. Примерами иллюстрируются методы и инструменты оценки точности цифровых моделей, интерактивной графической визуализации, сопоставления результатов.

ВВЕДЕНИЕ

Построение цифровых геологических, геоэкологических моделей является обязательной составляющей экспертных заключений в ряде сфер деятельности ([1] – [3]). Современная геологическая модель содержит обобщенное описание состава, структуры, формы объектов, состояния изучаемого фрагмента земной коры на разных стадиях. Она включает описание геологического строения, цифровые характеристики, обеспечивающие получение и графическую визуализацию объемного распределения структурно-вещественных комплексов и компонент модели. Соответствующее программное обеспечение компьютерных геологических моделей включает загрузку из разных источников и предобработку данных, корреляцию, формирование цифровых кубов фильтрационно-емкостных свойств, интерактивный анализ данных, визуализацию с применением графики разных типов, картопостроение.

Создание цифровых геологических моделей является относительно молодым направлением. Оно активно развивается. Одной из важнейших составляющих при этом является оценка адекватности и точности предлагаемых цифровых моделей, ключевыми являются вопросы автоматизации настройки, адаптации моделей, их сопоставления и оценки точности.

Моделирование объектов геологии предполагает использование сложных математических методов; программирование, реализация алгоритмов построения и адаптации геологических, геоэкологических моделей трудоемко; для поддержания подобных программных продуктов в актуальном состоянии, отвечающем постоянно возрастающим аппаратным возможностям, нужен большой коллектив опытных программистов ([2], [3]). Учитывая текущие возможности по обеспеченности системами компьютерной математики, более эффективным будет подход, основанный на интеграции в единый программный комплекс модулей современных версий систем компьютерной алгебры (СКА) и географических

информационных систем (ГИС). При этом следует понимать, что для решения задач обработки исходных данных, моделирования никакая конкретная ГИС не является полным набором пространственно-аналитических методов и средств анализа. Во многих случаях приходится комбинировать инструменты ГИС с программами статистического анализа данных, средствами для математически сложных вычислений, алгоритмов анализа и интерпретации пространственных данных.

I. ОСНОВНЫЕ ПОЛОЖЕНИЯ

В [2], [3] описаны предложенные и реализованные методические и технические решения, программные компоненты, которые включены в разработанный интегрированный компьютерный комплекс «Генератор геологической модели залежи» – ГГМЗ. Назначение комплекса тестирование, оценки точности настраиваемых геологических моделей на основе применения СКА, ГИС, «умных» методов адаптации моделей в процессе их эксплуатации, «самонастройки» моделей с учетом дополняемых данных фактического развития процессов. Разработка и функциональное наполнение комплекса ГГМЗ осуществлены с использованием системы компьютерной алгебры Wolfram Mathematica ([4], [5]) и геоинформационной системы Golden Software Surfer. При программировании в системе Mathematica модулей графики реализованы технические решения, описанные в [6] – [8]. Предусмотрена возможность, когда программный комплекс в конкретной конфигурации может эксплуатироваться после сборки и сохранения в формате вычисляемых документов CDF [9]. Расчеты и работа пользователя с CDF версией приложения возможны на любом персональном компьютере. При просмотре CDF версии, размещенной на вебсервере, программа просмотра автоматически подгружается в виде плагина браузера. Автономная работа с ПК возможна после инсталляции свободно распространяемого компонента CDF Player. Варианты дополнительных настроек, обеспечиваю-

щих интерактивность CDF версии, изложены в [7], [8].

Следует понимать, что создание и сопровождение геологической модели не предполагают нахождения единственного решения некоторой математической задачи. Субъективное мнение, квалификация эксперта – факторы, всегда присутствующие в подобной деятельности [1]. При создании моделей приходится оперировать с данными, точность которых различна, а некоторые исходные данные вовсе противоречивы; на разных участках плотность данных с замерами значительно различается. Поэтому для построения цифровых моделей важно иметь большой набор инструментов интерактивной обработки данных, имитации возможных ситуаций получения и коррекции входной информации, модули математической обработки данных. Система Mathematica обеспечивает пользователей широким набором встроенных функций численного анализа [10], методов и функций статистического анализа [11]. Требуемые функции подключаются в ГГМЗ простыми обращениями. Все этапы работы с данными в комплексе включают разнообразные варианты графической визуализации, протоколирование и сопоставление получаемых или помещенных в архив результатов [12], [7], [8]. Инструменты комплекса обеспечивают пользователя возможностями «поиграть» исходными данными и сопоставить результаты с подготовленными эталонами, причем, с возможностями импорта и экспорта данных и изображений, масштабирования. Широкие возможности обмена данными важны для одновременной работы в нескольких программных средах.

II. МЕТОДЫ И ИНСТРУМЕНТЫ АНАЛИЗА ПОГРЕШНОСТЕЙ

В комплексе ГГМЗ для визуализации моделей и их сопоставления реализованы более десяти вариантов 1D, 2D и 3D графики, в том числе модули получения и оформления карт и вставок на них, графиков на профилях, 3D визуализации возможных разрезов (вертикальные, горизонтальные), простых и сложных отсечений. Основные возможности программных модулей визуализации для нескольких типичных примеров описаны в упомянутых выше статьях [2], [3]. В настоящей работе в дополнение иллюстрациям, которые дают только общее представление об отличиях форм поверхностей, описаны инструменты получения и визуализации на формируемых картах невязки различий рассчитываемых разными способами или по отличающимся исходным данным восстанавливаемых пространственных цифровых полей. Из карт распределения невязки хорошо видно, какие участки воспроизводятся с превышением, какие с понижением. Также описаны инструменты получения числовых значений отклонений для интегральных оценок по выделяемым областям и участкам.

ЗАКЛЮЧЕНИЕ

В докладе будут изложены и проиллюстрированы примерами предложенные и обоснованные методические и технические решения, рекомендации анализа погрешностей создаваемых и сопровождаемых цифровых геологических моделей. Акцентируется внимание на необходимости при компьютерном моделировании и оценке его качества, как визуального, так и числового сопоставления результатов, расчета и выявления интегральных и локальных характеристик ошибок.

III. СПИСОК ЛИТЕРАТУРЫ

1. Халимов, Э. М. Детальные геологические модели и трехмерное моделирование, Нефтегазовая геология. Теория и практика / Э. М. Халимов // Нефтегазовая геология. Теория и практика. –2012. –№ 3 (7). –С. 1–10.
2. Taranchuk, V. B. The integrated computer complex of an estimation and adapting of digital geological models / V. B. Taranchuk // Studia i Materialy. Volume 14. –2017. –№ 2. –С. 73–86 (in Engl.)
3. Таранчук, В. Б. Интегрированный программный комплекс тестировщика геологических моделей / В. Б. Таранчук // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. –2017. –№ 16 (265). –С. 148–159.
4. Таранчук, В. Б. Введение в язык Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 51 с.
5. Таранчук, В. Б. Основы программирования на языке Wolfram : учеб. материалы для студентов фак. прикладной математики и информатики спец. 1-31 03 04 «Информатика» / В. Б. Таранчук // Минск: БГУ, 2015. – 49 с.
6. Таранчук, В. Б. Особенности функционального программирования интерактивных графических приложений / В. Б. Таранчук // Вестник Самарского государственного университета. Естественнонаучная серия, раздел Математика: –2015. –№ 6 (128). –С. 178–189.
7. Таранчук, В. Б. Возможности и средства Wolfram Mathematica для разработки интеллектуальных обучающих систем / В. Б. Таранчук // Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика: –2015. –№ 1 (198) том 33. –С. 102–110.
8. Taranchuk, V. B. Methods and tools of development of interactive teaching materials on the computer mechanics / V. B. Taranchuk, M. A. Zhuravkov // Studia i Materialy. –2016. –№ 11. –С. 69–83 (in Engl.).
9. Таранчук, В. Б. О подготовке и распространении на базе системы Mathematica интерактивных графических приложений / В. Б. Таранчук, В. А. Куликович // Информатизация образования: –2015. – № 1 (75). –С. 3–13.
10. Морозов, А. А. Программирование задач численного анализа в системе Mathematica : учеб. пособие / А. А. Морозов, В. Б. Таранчук // Минск : БГПУ – 2015. – 145 с.
11. Таранчук, В. Б. О возможностях и инструментах реализации вероятностно-статистических методов в системе Mathematica / В. Б. Таранчук // Теория вероятностей, случайные процессы, математическая статистика и приложения : материалы Междунар. науч. конф., посвящ. 80-летию проф., д-ра физ.-мат. наук Г.А. Медведева, Минск : РИВШ, 2015. –С. 320–325.
12. Таранчук, В. Б. Графический сервис вычислительного эксперимента : учеб.-метод. пособие / В. Б. Таранчук // Минск : БГУ –2009. – 124 с.

О РАЗРАБОТКЕ ПРИЛОЖЕНИЯ ДЛЯ ОПТИМИЗАЦИИ РАБОТЫ ЛОГИСТИЧЕСКИХ КОМПАНИЙ НА БАЗЕ МОБИЛЬНОЙ ПЛАТФОРМЫ IOS

Ткачев И. Е., Рудикова Л. В.

Кафедра современных технологий программирования, Гродненский государственный университет имени Янки Купалы

Гродно, Республика Беларусь

E-mail: tkachev_ie_15@mf.grsu.by, rudikowa@gmail.com

Излагаются общие подходы к реализации мобильного приложения на платформе iOS, которое предназначено для оптимизации рабочих процессов в логистических компаниях и упрощения коммуникации между сотрудниками. Приводится основная функциональность программного обеспечения, связанного с разработкой указанного мобильного приложения.

ВВЕДЕНИЕ

В настоящее время сфера высоко интегрированных приложений для портативных компьютеров, смартфонов и сотовых телефонов развивается достаточно интенсивно. Однако, актуальными являются разработки для поддержки задач в специализированных предметных областях. Так, например, логистические компании, занимающиеся доставкой грузов наземным путем. В настоящее время спрос на логистические услуги постоянно возрастает. Это вызвано ростом экономики и усилением конкуренции среди производителей. Проблема конкурентоспособности решается путем совершенствования процедур складирования, транспортировки, распределения товаров между потребителями. Перед логистами встает задача регулировки и контроля доставки каждого груза, ведь, все поставки должны быть выполнены в срок и без непредвиденных ситуаций. Кроме того, должна поддерживаться постоянная связь с водителями, которых в компании может насчитываться более сотни. Все эти факты делают невозможным нормальное функционирование, централизованный контроль и развитие компании.

I. ЭТАПЫ РАЗРАБОТКИ ПРИЛОЖЕНИЯ

Работу над реализацией приложения «Delievery Assistant» можно разбить на следующие этапы: проектирование веб-сервиса для хранения и обновления информации, связанной с доставкой, а также обеспечивающего связь между пользователями; проектирование базы данных; получение модели функций разрабатываемого приложения; разбиение приложения на модули в соответствии с выполняемыми функциями, проектирование интерфейса; реализация базы данных и логики взаимодействия с ней; реализация интерфейсов представлений; создание классов, используемых для привязки полученных данных к представлениям; реализация связи базы данных с интерфейсом. При первом запуске пользователю необходимо создать аккаунт, используя свою электронную почту, и прове-

сти авторизацию. Это главный аккаунт, который имеет возможность управлять системой и создавать подчиненные аккаунты соответствующего уровня. Далее приложение создает базу данных, а также необходимые таблицы, которые в будущем заполняются данными. В процессе работы приложения модифицируется содержимое базы данных, информация, полученная с веб-сервера, заносится в локальную базу данных. Это необходимо для доступа пользователя к собственным данным при отсутствии соединения с сетью. Следует отметить, что при добавлении данных пользователем в базу при отсутствии доступа к сети, в дальнейшем при наличии доступа в сеть, произойдет синхронизация данных с веб-сервером.

II. ОБЩИЕ ПОДХОДЫ К АРХИТЕКТУРЕ РЕАЛИЗАЦИИ ПРИЛОЖЕНИЯ

Рассмотрим основные компоненты архитектуры мобильного приложения для оптимизации работы логистических компаний. Приложение «Delievery Assistant» состоит из следующих частей: базы данных, в которой хранятся данные; веб-сервера, сохраняющего пользовательские данные; веб-сервера, обеспечивающего связь между пользователями; пользовательского интерфейса, предоставляющего пользователю доступ к функциям приложения; сервиса, предоставляющего GPS данные; авторизации через электронную почту. Для того чтобы пользователь мог использовать функционал приложения, ему необходимо совершить авторизацию через свою электронную почту. Авторизация реализована через технологию OAUTH 2.0, которую поддерживает SPRING BOOT (на его основе построен веб-сервер), что подразумевает генерацию token-авторизации. Далее уведомление о входе будет выслано на почту пользователя. Для начала работы приложение должно получить token пользователя, который хочет авторизоваться в приложении. В основе приложения лежит архитектурный принцип SOLID, поэтому приложение разбито на модули, каждый из которых выполняет свою функцию. При этом моду-

ли используют общие сервисы, которые внедряются в них извне, что позволяет ослабить зависимости между ними, и повысить возможность их дальнейшего переспользования. Работа приложения начинается с модуля Login, в ходе работы которого проверяется, был ли пользователь уже авторизован или нет. Если пользователь уже авторизовался, то в работу включается модуль Main и задействует первую вкладку приложения, которая также является отдельным модулем и называется Status. Пользователь получает доступ к основному функционалу приложения. Однако, если при обращении к UserDefaults (используется в приложении для создания именованных ассоциативных массивов типа «ключ – значение», которые могут быть использованы различными компонентами приложения) отсутствует token, то пользователю предлагается произвести авторизацию. Интерфейс авторизации довольно прост и понятен. Login модуль обращается к LoginHelper (класс, который инкапсулирует логику авторизации пользователя). Если же у пользователя логин отсутствует, ему будет предложено завести новый аккаунт и в случае подтверждения будет вызван модуль CreateAccount. После ввода комбинации логина-пароля, на почту пользователю будет выслано сообщение с кодом подтверждения. Если же авторизация прошла успешно, то полученный token из CreateAccount модуля передается в Main. Однако, если при авторизации произошла ошибка (т.е. после запроса авторизовать пользователя, веб-сервер вернул ошибку), то выводится сообщение об ошибке. После того как был получен token из Login или CreateAccount модуля, Main модуль обращается к UserDefaults для записи полученного token в хранилище типа «ключ-значение». Это необходимо для того, чтобы пользователь каждый раз не вводил свою комбинацию логин/пароль. Далее Main вызывает модуль Status для предоставления основного функционала приложения пользователю, либо предлагает настроить аккаунт, если аккаунт был только что создан. Main модуль является контейнером для других модулей и отображает их в виде вкладок. Такая архитектура упрощает синхронизацию данных между модулями. Само представление данного модуля наследуется от класса UITabBarController. Он содержит следующие вкладки: Status, Chat, Map, Settings. Каждая вкладка – это самостоятельный модуль, который выполняет строго свою функцию и построен на основе VIPER. После того как Status-модуль был вызван, он обращается к NetworkService (общий для всех модулей сервис для работы с сетью) для загрузки данных. Далее NetworkService проверяет наличие Интернета и, если он отсутствует, то все данные читаются из локальной базы данных, а, если соединение присутствует, то NetworkService обращается к веб-

серверу для получения данных. Status модуль отображает всю необходимую информацию о текущих рейсах. Chat модуль реализует чат с другими сотрудниками организации. Map отображает текущее расположение грузов (использует Google Maps API), Setting позволяет осуществлять настройки аккаунта. NetworkService обращается к веб-серверу, где в качестве идентификации пользователя используется User-id, сгенерированный на основе адреса электронной почты. Полученные данные хранятся в формате JSON, поэтому для удобной работы с ними используется стандартная библиотека которая структурирует данные для удобного пользования. После загрузки данных NetworkService обращается к DataBaseManager для того, чтобы записать полученные с веб-сервера данные в локальную базу данных. Это необходимо для использования приложения без соединения с Интернетом. DataBaseManager взаимодействует с базой данных SQLite используя встроенный инструмент CoreData, который упрощает работу с базой данных. После сохранения данных в базу данных, полученная информация передается непосредственно в модуль. Для загрузок картинок используется класс ImageLoader, инкапсулирующий логику загрузки и кэширования картинок. Модуль передает в ImageLoader адрес картинки в сети и UIImageView, в который необходимо загрузить файл картинки. ImageLoader использует NSCache для хранения картинок в оперативной памяти устройства. После того как был передан URL и UIImageView, загрузчик картинок, прежде всего, проверяет наличие файла картинки в кэше: если картинка отсутствует, то загрузчик, используя HttpClient, загружает ее. После загрузки картинка помещается в кэш. Для того, чтобы картинка не загружалась дважды и при следующем обращении к картинке, она уже не будет загружаться из сети, а будет извлекаться из кэша, что обеспечивает экономию трафика. После загрузки картинка присваивается указанному UIImageView.

III. ЗАКЛЮЧЕНИЕ

Таким образом, разработанное приложение предназначено для использования в логистических компаниях, которые заинтересованы в оптимизации их внутренней работы и контроля рабочих процессов. Приложение «Delievery Assistant» предоставляет возможность создания и настройки списка текущих доставок, отслеживания доставляемых грузов и общения между сотрудниками компании, вести централизованный учет расходов. В дальнейшем приложение может быть адаптировано и использоваться для связи грузоотправителя и грузоперевозчика, заключения между ними контракта, ведение учета на складах.

ИСПОЛЬЗОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ПРИ РАСПОЗНАВАНИИ И АНАЛИЗЕ СОСТОЯНИЙ ДИНАМИЧЕСКИХ ОБЪЕКТОВ

Тюшкевич М. В.

Кафедра информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: markt93@mail.ru

В данной статье рассмотрены подходы к решению задач распознавания динамических объектов и анализ их состояний. Дополнительной сложностью в распознавания динамических объектов является то, что помимо статических признаков объекта существуют и динамические характеристики. При решении задачи следует так же следить за изменением пространственных данных во времени. В качестве решение было предложено использование рекуррентных нейронных сетей. Для обучения сети применяется алгоритм обратного распространения ошибки.

ВВЕДЕНИЕ

В настоящее время широко развиваются системы распознавание динамических объектов, которые нашли применение в различных отраслях. Для построения современных систем определения динамических объектов характерны следующие основные проблемы и задачи: обнаружение объекта интереса на сложном изменяющемся фоне; измерение пространственных параметров и оценка динамических параметров объектов по наборам последовательных кадров; сопровождение объектов; классификация и распознавание динамических объектов. Для решения такого рода задач широкое применение получили искусственные нейронные сети.

В задаче распознавания динамических объектов входной сигнал, подлежащий распознаванию, рассматривается не в статике, а в динамике и основная информация об образе содержится в траектории изменения входного сигнала во времени. В результате образ получает новое измерение - время, что значительно усложняет решение задачи распознавания. Анализ пространственно-временных данных позволяет выявлять не только статические, но и динамические признаки объектов наблюдения. В этом случае задачу распознавания можно определить как классификацию совокупностей состояний или как классификацию траекторий, решение которой не может быть найдено классическими методами распознавания, т.к. временные переходы могут порождать преобразования изображений, не описываемые известными аналитическими зависимостями[1].

Также наряду с задачей распознавания динамических объектов возникают задачи распознавания активных действий объектов и событий сцены, например, для выявления несанкционированных действий в местах скопления людей или определении жанра сцены для индексации в мультимедийных базах данных. Если рассматривать задачу распознавания объектов и собы-

тий по последовательностям изображений в виде единого процесса, то наиболее целесообразным является системный иерархический подход с элементами параллельной обработки на каждом уровне[2].

1. РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ КАК ДИНАМИЧЕСКАЯ СИСТЕМА

Решение задачи идентификации динамического объекта является одной из ключевых задач при построении нейросетевой системы управления. В качестве идентификатора используются методы теории нейронных сетей. Применительно к динамическим объектам сеть должна обладать архитектурой, позволяющей аппроксимировать временные последовательности. Из теории нейронных сетей известна теорема об универсальной аппроксимации многослойной нейронной сети, согласно которой достаточно применения сети с одним скрытым слоем для построения равномерной аппроксимации с точностью для любого обучающего множества, которое может быть задано входами и желаемыми откликами системы[3]. Дальнейшая работа по созданию модели идентификации динамического объекта опирается на вышеприведенную теорему.

Выбор архитектуры нейронной сети сводится к выбору среди рекуррентных многослойных сетей. Подобные сети можно получить из классической многослойной нейронной сети путем введения обратных связей (см. рис. 1), благодаря которым они имеют собственную динамику, что и отличает их от статических НС. Структура рекуррентной нейронной сети является её основной характеристикой, от которой напрямую зависят все её свойства. Обратные связи могут быть от выхода сети ко входу и от внутреннего слоя сети ко входу. При этом объект может быть рассмотрен в пространстве состояний либо, при отсутствии информации о внутренней структуре, как модель «вход-выход».

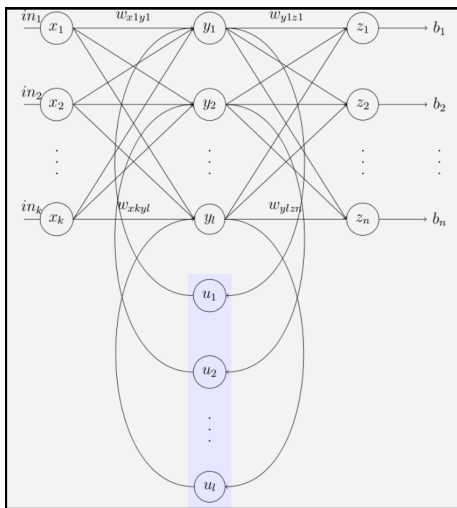


Рис. 1 – Пример простой рекуррентной сети

Необходимым и достаточным условием для устойчивости сети с обратной связью будет наличие единичной задержки в обратной связи. РНС может быть устойчива (не устойчива) в малом и в большом. Устойчивость в большом гарантирует устойчивость во всем пространстве состояний, а устойчивость в малом – лишь в определенных точках. Для устойчивости в большом необходимо подобрать специальную функцию, удовлетворяющую условиям второй теоремы Ляпунова (об асимптотической устойчивости). Для устойчивости в малом необходимо произвести линеаризацию системы в точке равновесия и проверить в ней устойчивость.

II. ОБУЧЕНИЕ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ

Для обучения РНС можно выделить несколько различных по своей сути групп подходов:

- использование известных соотношений между параметрами РНС и её динамикой;
- использование алгоритмов обучения с учителем на базе методов оптимизации по типу алгоритма обратного распространения ошибки;
- использование алгоритма обучения с учителем;
- использование алгоритмов обучения без учителя;
- отсутствие обучения в классическом смысле изменения параметров системы (неявное обучение) – настройка весовых коэффициентов случайными значениями.

Наиболее используемый в настоящее время алгоритм – обучение с учителем на базе методов оп-

тимизации по типу алгоритма обратного распространения ошибки [4]. Он позволяет с достаточной для практического применения точностью оптимизировать веса нейронной сети. В его основе – метод градиентного спуска, когда веса корректируются в соответствии с направлением уменьшения градиента общей для всего обучающего множества ошибки. Он состоит из нескольких основных шагов: инициализации, представления обучающей выборки, прямого прохода расчета сети, обратного прохода и вычисления текущей ошибки.

При инициализации весов используется генерация случайных чисел. После инициализации сети поочередно предъявляются все обучающие пары «вход-выход», предварительно снятые с объекта, подвергаемого идентификации при различных условиях. Для формирования выборки наличие большего количества примеров повышает качество будущей идентификации.

III. ЗАКЛЮЧЕНИЕ

Существует много разновидностей, решений и конструктивных элементов рекуррентных нейронных сетей.

Трудность рекуррентной сети заключается в том, что если учитывать каждый шаг времени, то становится необходимым для каждого шага времени создавать свой слой нейронов, что вызывает серьезные вычислительные сложности. Кроме того многослойные реализации оказываются вычислительно неустойчивыми, так как в них как правило исчезают или сильно перевешивают веса. Если ограничить расчёт фиксированным временным окном, то полученные модели не будут отражать долгосрочных трендов. Различные подходы пытаются усовершенствовать модель исторической памяти и механизм запоминания и забывания.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Дуда, Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт // М.: изд-во «Мир», 1978. – 512 с.
2. Малинецкий, Г. Г. Нелинейная динамика и хаос. Основные понятия: Учеб. пособие / Г. Г. Малинецкий, А. Б. Потапов // М.: Ком-Книга, 2006. – 240 с.
3. Хайкин, С. Нейронные сети: полн. курс / С. Хайкин; Пер. с англ. – 2-е изд., // – М. – СПб. – Киев: ИД «Вильямс», 2006. – 1103 с.
4. Никитин, К. В. Новый подход к применению рекуррентных нейронных сетей для решения задач распознавания образов / К. В. Никитин, Е. Н. Бендерская // Научно-технические ведомости СПбГПУ. – СПб.: Изд-во Политехн. ун-та, – 2007. – № 4-2 (52). – С. 85–92.

АНАЛИЗ АЛГОРИТМОВ ВЫБОРА ПРИЗНАКОВ ЭКЗОНОВ

Волков А. В., Яцков Н. Н., Гринев В. В.

Кафедра системного анализа и компьютерного моделирования, кафедра генетики, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: andrei@cybergizer.com, yatskou@bsu.by, grinev@bsu.by

Рассмотрена задача сокращения размерности пространства признаков экзонов человека с целью увеличения точности определения их генной принадлежности. Выполнен сравнительный анализ алгоритмов отбора признаков при варьировании алгоритмов индуктивного обучения.

ВВЕДЕНИЕ

Исследование организации и функционирования генов человека является важной задачей биоинформатики [1]. Гены состоят из экзонов и интронов. Экзон характеризуется большим количеством признаков (более 1000), в то же время число экзонов, принадлежащих гену, невелико (как правило, менее 200). Проблема большого числа признаков и относительно малого числа объектов наблюдений характерна для всей области биоинформатики в целом в связи с трудностями для алгоритмов индуктивного обучения.

Использование алгоритмов автоматического отбора признаков позволяет снизить размерность решаемой задачи [2]. Эффективность упомянутых алгоритмов определяется свойствами исследуемых наборов данных и числом объектов наблюдения [3], что формирует необходимость исследования алгоритмов отбора признаков в контексте экзонных данных. Так в настоящей работе исследованы алгоритмы отбора признаков семейства Relief [5].

Целью работы является выяснение принципиальной возможности предсказания генной принадлежности экзонов, а также исследование эффективности алгоритмов семейства Relief в задачах отбора признаков экзонов.

1. МЕТОДОЛОГИЯ

Экспериментальные данные получены из базы данных Ensembl [4] и содержат 1762 уникальных экзона, принадлежащих 14 произвольно отобраным генам. Каждый экзон дополнительно охарактеризован с помощью 1198 численных признаков: 429 признаков непосредственно самих экзонов и 769 признаков фланкирующих участков нуклеотидных последовательностей (длина цепи составляет 100 нуклеотидов).

Среди алгоритмов отбора признаков широкое распространение получили методы-фильтры [2], что обусловлено простой структурой, вычислительной эффективностью и независимостью от типа используемого алгоритма индуктивного обучения. Абсолютное большинство методов-фильтров являются унивариативными [2]. Исключение составляет семейство алгорит-

мов на основе метода Relief [5], способные учитывать зависимости между признаками. Алгоритмы SURF [5], MultiSURF [5], SURFStar [5] являются наиболее популярными Relief-алгоритмами в приложениях биоинформатики и включены в состав фреймворка ReBATE (англ. Relief-Based Algorithm Training Environment) [5].

В работе рассмотрены три алгоритма индуктивного обучения: метод k -ближайших соседей [6], машина опорных векторов с линейным ядром [7] и наивный байесовский классификатор [8]. Выбор алгоритмов обусловлен их популярностью, разнородностью подходов к индуктивному обучению, отсутствием встроенных механизмов отбора признаков.

В работе использовано скоринговое правило Брайера [9] для оценки качества вероятностного прогноза алгоритма индуктивного обучения.

Блок-схема организации вычислительного эксперимента представлена на рис. 1.

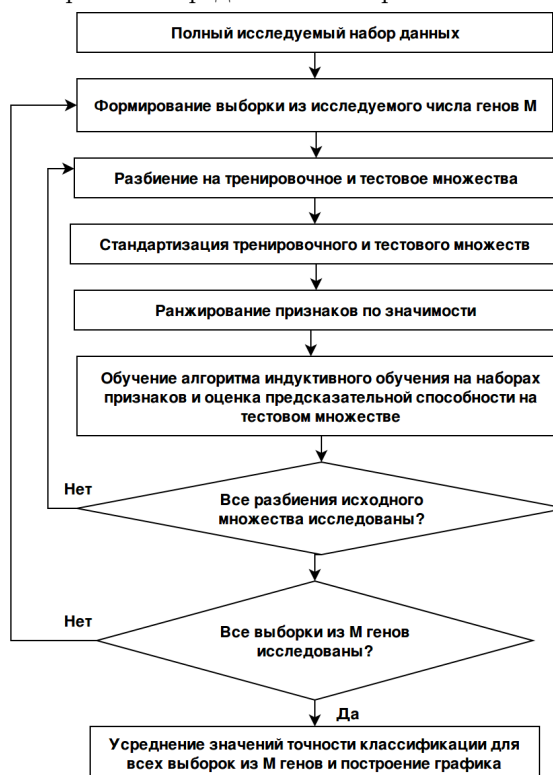


Рис. 1 – Блок-схема организации вычислительного эксперимента

Представленный подход позволяет исследовать зависимость оценки предсказательной способности алгоритмов индуктивного обучения от числа ранжированных по информативности признаков.

II. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Исследована эффективность алгоритмов отбора признаков в задачах предсказания генной принадлежности экзонов человека при варьировании алгоритмов индуктивного обучения. Установлен факт значимой разделимости между экзонами, принадлежащими различным генам. Пример зависимости представлен на рис. 2.

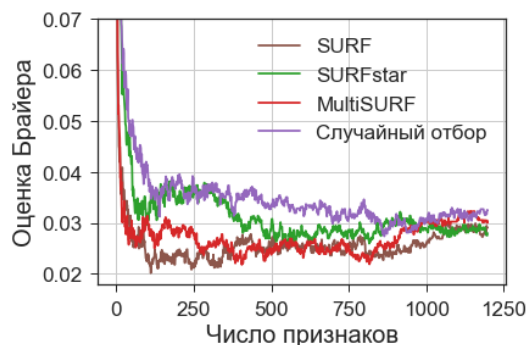


Рис. 2 – Зависимость оценки по скоринговому правилу Брайера от числа ранжированных по информативности признаков (2 гена, метод 1 ближайшего соседа)

Тренировка алгоритмов индуктивного обучения на признаках фланкирующих интронов обеспечивает более высокую предсказательную способность классификаторов по сравнению с тренировкой алгоритмов индуктивного обучения на признаках экзонных нуклеотидных последовательностей (рис. 3). Это наблюдение само по себе представляет большой интерес и требует дальнейшего детального изучения с помощью методов биоинформатики, а также экспериментальных методов молекулярной биологии.

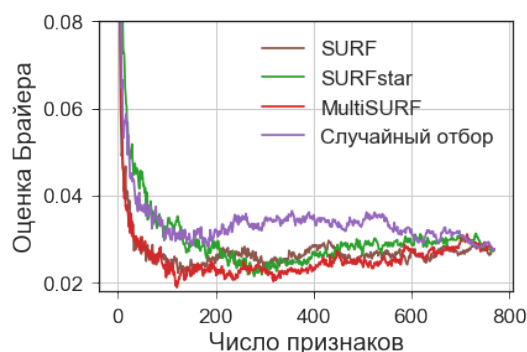


Рис. 3 – Зависимости оценок по скоринговому правилу Брайера для ранжированного ряда признаков фланкирующих интронов

III. ЗАКЛЮЧЕНИЕ

Установлен факт значимой разделимости между экзонами, принадлежащими разным генам. Показано, что использование алгоритмов автоматического отбора в сочетании с методом k-ближайших соседей (1 ближайший сосед) позволяет уже на 15 признаках достигать 96% точности предсказания генной принадлежности, что на 3.6% выше, чем случайный отбор признаков и значительно вычислительно эффективнее анализа полного набора из 1198 признаков. Обнаружено, что более низкие значения счетов Брайера (0,02) соответствуют обучению алгоритмов на признаках фланкирующих интронов, в сравнении с величиной 0,07 для признаков экзонных нуклеотидных последовательностей (метод одного ближайшего соседа в сочетании с алгоритмом MultiSURF).

IV. СПИСОК ЛИТЕРАТУРЫ

1. Grinev V. V., Migas A. A., Kirsanova A. D., Mishkova O. A., Siomava N., Ramanouskaya N. V., Vaitsiankova A. V., Ilyushonak I. M., Nazarov P. V., Vallar L., Aleinikova O. V. Decoding of exon splicing patterns in the human RUNX1-RUNX1T1 fusion gene // *Int. J. Biochem. Cell Biol.* 2015. Vol. 68. P. 48-58
2. Li JD, Cheng KW, Wang SH, Morstatter F, Trevino RP, Tang JL, Liu H (2016) Feature selection: a data perspective, vol 3, pp 1-73. arXiv:1601.07996
3. Oreski D, Oreski S, Klicek B. Effects of dataset characteristics on the performance of feature selection techniques. *Appl Soft Comput.* 2017;52:109119.
4. Aken, B. L. The Ensembl gene annotation system. B.L. Aken, S. Ayling, D. Barrell, L. Clarke, V. Curwen, S. Fairley, J. Fernandez Banet, K. Billis, C. Garcia Giron, T. Hourlier, et al. (2016) Database (Oxford), doi: 10.1093/database/baw093
5. Urbanowicz, R. J., Olson, R. S., Schmitt, P., Meeker, M., Moore, J. H., 2018. Benchmarking relief-based feature selection methods for bioinformatics data mining. *Journal of biomedical informatics.*
6. Altman, N. S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression". *The American Statistician.* 46 (3): 175-185.
7. Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks". *Machine Learning.* 20 (3): 273-297.
8. Russell, Stuart; Norvig, Peter (2003) [1995]. *Artificial Intelligence: A Modern Approach* (2nd ed.). Prentice Hall. ISBN 978-0137903955.
9. Gneiting, T. and Raftery, A. E. (2007) Strictly proper scoring rules, prediction, and estimation. *J. Am. Statist. Ass.*, 102, 359-378.

ИМИТАЦИОННАЯ МОДЕЛЬ ДЛЯ ТЕСТИРОВАНИЯ АЛГОРИТМОВ ОТБОРА ПРИЗНАКОВ

Волков А. В., Яцков Н. Н., Гринев В. В.

Кафедра системного анализа и компьютерного моделирования, кафедра генетики, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: andrei@cybergizer.com, yatskou@bsu.by, grinev@bsu.by

Предложен алгоритм имитационного моделирования для генерации многомерных наборов данных, учитывающий информативность групп признаков. Представлены результаты исследования алгоритмов отбора признаков на смоделированных данных.

ВВЕДЕНИЕ

Исследование алгоритмов отбора признаков [1] на тестовых наборах данных позволяет определять достоинства и недостатки алгоритмов. В настоящее время в научном сообществе наиболее популярным является тестирование существующих и разрабатываемых алгоритмов отбора признаков на реальных наборах данных. В качестве примера следует выделить репозиторий тестовых наборов данных [2] для различных предметных областей, включающих анализ текстовой информации, изображений и биологических данных. Недостатком данного подхода является невозможность апробирования алгоритмов в широком спектре экспериментальных условий. Кроме того, оптимальный набор признаков для подобного рода наборов данных зачастую является неизвестным. Задачу выбора наиболее подходящего алгоритма можно упростить, используя имитационное моделирование и синтетические данные [3,4]. Преимуществами данного подхода являются:

1. возможность варьирования параметров исследуемых наборов данных (таких как число информативных, шумовых и избыточных признаков, число объектов наблюдения, задание уровня шума для меток классов, регулирование степени информативности признаков) облегчает получение обоснованных выводов по результатам проведенного анализа;
2. априорное знание оптимального набора признаков позволяет осуществлять надежную и автоматизированную оценку степени близости полученного результата отбора признаков к истинному.

Существующие имитационные модели генерации кластеров многомерных данных имеют ряд ограничений [5], среди которых наиболее существенным является невозможность явной задачи степени информативности признака или групп признаков.

Цель работы – разработка имитационной модели учитывающей информативность признаков объектов данных.

I. МЕТОДОЛОГИЯ

Входные параметры имитационной модели: число кластеров в генерируемом наборе данных; число групп признаков с заданной степенью делимости между кластерами; число шумовых признаков; число избыточных признаков; уровень шума в метках классов; размеры кластеров. В рамках разработанной модели доступны два алгоритма генерации кластеров данных с заданной степенью делимости:

1. на основе размещения кластеров данных в вершинах гиперкуба, где величина длины ребра определяется заданным индексом делимости [6].
2. на базе вычисления метрики свободного пространства между двумя произвольными кластерами [7,8];

Алгоритм осуществляет генерацию гауссовых кластеров со случайными ковариационными матрицами, определяющими форму кластеров.

Пример сгенерированного набора данных из трех кластеров представлен на рис. 1.

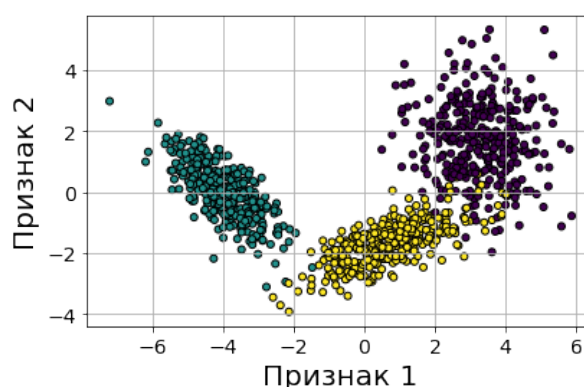


Рис. 1 – Пример сгенерированного набора данных из трех кластеров

Число атрибутов в рамках выделенной группы признаков является меньшим либо равным числу кластеров минус единица [9].

Блок схема алгоритма имитационного моделирования представлена на рис. 2.

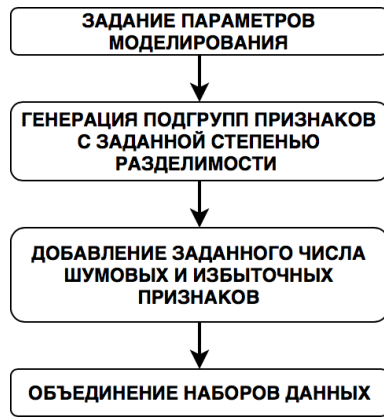


Рис. 2 – Блок схема алгоритма имитационного моделирования

Для проверки применимости предложенного подхода выбраны релевантные представители фильтрующих алгоритмов отбора признаков: алгоритм счета Фишера [10] и алгоритм ReliefF [11].

II. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Исследована применимость предложенного подхода имитационного моделирования. В ходе анализа сгенерированы наборы данных с различными степенями разделимости. Работоспособность разработанного алгоритма имитационного моделирования подтверждена на примерах сравнительного анализа алгоритмов отбора признаков. На рис.3 представлен пример значений счета Фишера для набора параметров:

1. подход на основе гиперкуба для генерации трех групп признаков со значениями индекса разделимости 1.2, 0.8 и 0.3 соответственно;
2. 300 наблюдений в каждом кластере;
3. отсутствие дополнительных избыточных и шумовых признаков, а также шума в метках классов;
4. 15 классов;
5. 14 признаков в каждой группе;

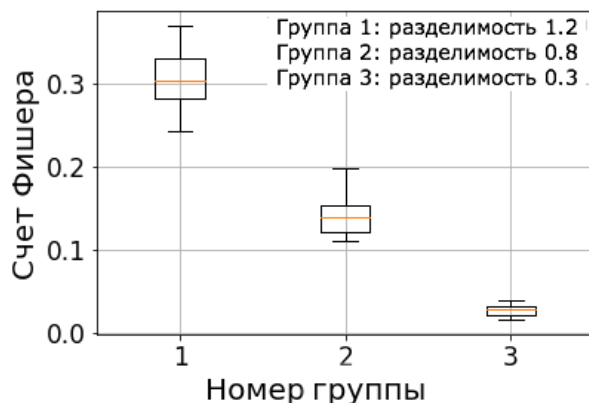


Рис. 3 – Пример диаграммы значений счета Фишера

Наличие статистически значимого различия средних значений счета Фишера для каждой из групп в приведенном примере подтверждается результатом однофакторного дисперсионного анализа: значение р-величины составляет 7.5×10^{-23} .

III. ЗАКЛЮЧЕНИЕ

В настоящей работе предложена и исследована имитационная модель для задач тестирования алгоритмов отбора признаков. Уникальной особенностью модели является генерация многомерных наборов данных с заданной степенью разделимости в рамках групп признаков. Результаты исследования алгоритмов отбора признаков на смоделированных данных демонстрируют адекватность разработанной модели.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Robert P Trevino, Jiliang Tang, and Huan Liu. Feature selection: A data perspective. arXiv preprint arXiv:1601.07996, 2016
2. Feature selection data set repository [Electronic resource] / Jundong Li et al. // Arizona State University. Mode of access: <http://featureselection.asu.edu/datasets.php>. Date of access: 25.09.2018.
3. V. Bolon-Canedo, N. Sanchez-Marono, A. Alonso-Betanzos, "An ensemble of filters and classifiers for microarray data classification," *Journals Pattern Recognition*, vol. 45, pp. 531-539, 2012.
4. Belanche L.A., Gonzalez FF, Review and evaluation of feature selection algorithms in synthetic problems. <http://arxiv.org/abs/1101.2320>
5. P. Ghosh, Arka, Ranjan Maitra and D. Anna Peterson, "A Separability Index for Distance-based Clustering and Classification Algorithms", 2012.
6. I. Guyon, "Design of experiments for the NIPS 2003 variable selection benchmark", 2003.
7. W. Qiu and H. Joe. Generation of random clusters with specified degree of separation. *Journal of Classification*, 23:315–334, 2006.
8. W. Qiu and H. Joe. Separation index and partial membership for clustering. *Computational Statistics and Data Analysis*, 50:585–603, 2006.
9. McLachlan, G. J. (2004). *Discriminant Analysis and Statistical Pattern Recognition*. Wiley Interscience. ISBN 0-471-69115-1.
10. P. E. H. R. O. Duda and D. G. Stork. *Pattern Classification*. Wiley-Interscience Publication, 2001.
11. M. Robnik-Sikonja and I. Kononenko. Theoretical and empirical analysis of relief and rrelief. *Machine Learning*, 53(1-2):23–69, 2003.

ИДЕНТИФИКАЦИЯ УЧАСТНИКОВ ОБЩЕСТВЕННЫХ МЕРОПРИЯТИЙ НА ОСНОВЕ ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ

Якубович А.В., Лежнёв А.В.

Кафедра информационных технологий автоматизированных систем, кафедра систем управления,
Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: yakubovichandreiv@gmail.com, anton.lezhnev@yandex.by

В статье предлагается анализ данных пользователей социальных сетей для сбора информации о реальных событиях на основе определения местоположения, схожести контента и социальных связей источников. Рассматривается регрессия гауссовского процесса для определения расстояния между источником сообщения о событии и предполагаемым событием.

ВВЕДЕНИЕ

Широкое использование социальных сетей, особенно их услуг на основе определения местоположения, превратило социальные сети в важный источник информации о реальных событиях. В последние годы было разработано множество систем обнаружения событий, использующие геотегированные посты из социальных сетей. Данная технология востребована правоохранительными чиновниками, полицией, и средствами массовой информации так как позволяет определить участников мероприятия, от которых можно получить ценную информацию о самом событии или расследовать причастность подозреваемых если событие связано с преступлением или терроризмом. Однако из-за высокой неопределенности мобильности людей, и низкой вероятности того, что пользователи делятся своей информацией о местоположении, трудно идентифицировать пользователей на месте события.

В этой статье мы предлагаем модель на основе функции сплавленной регрессии гауссовского процесса, которая использует три влиятельных фактора в социальных сетях для идентификации источников сообщения на месте: влияние мобильности, сходство контента и социальные отношения источников. Объединив эти факторы, мы можем оценить расстояние между источником сообщения и событием даже если профиль местоположения источника неизвестен [1].

I. ПОСТАНОВКА ЗАДАЧИ

Для каждого социального события E , которое происходит в окне времени T , обычно существует группа связанных с событиями источников $U_{all}(u_1, u_2, \dots)$, которые размещают соответствующую информацию о событии, включая активные источники U_{act} и обычные источники U_{nor} . Однако сведения о местоположении могут быть неопределёнными, учитывая конфиденциальность участников. Поэтому все абсолютные данные по GPS будут преобразованы в относительные расстояния как в моделировании при-

знаков, так и в вычете результатов. И расстояние между каждым источником U_i в U_{all} и событием E определяется как $Dis(U_i, E)$. Задача исследования - оценка $Dis(U_i, E)$ для содействия выявлению пользователей на местах. Предполагается, что малое $Dis(U_i, E)$ означает более высокую возможность что U_i присутствует на месте. Поэтому, основная проблема - как определить $Dis(U_i, E)$ точно и эффективно.

В частности, социальное событие E может быть представлен как совокупность трёх величин $E < \Omega, \Theta, \Phi >$, где Ω - множество источников, Θ тема и Φ картина социальных связей извлеченная из Ω . Каждый источник U_i в U_{all} также моделируется тремя параметрами $U < \omega, \theta, \phi >$, где ω - недавние упоминания в сообщениях источника, которые указывают на его/её социальные связи, θ содержит историю тем поднимаемых источником и ϕ представляет индивидуальную картину мобильности этого источника. В соответствии с этими двумя определениями, извлекаются три параметра: Влияние мобильности, Сходство содержания и Социальные отношения (F_{MI}, F_{CS}, F_{SR}) между источником и событием. Предположим, что существует функция подобия $Sim(x, y)$, которая измеряет сходство между двумя данными формата x и y .

Можно предположить, что $F_{MI} \propto Sim(\Phi, \phi)$, $F_{CS} \propto Sim(\Theta, \theta)$ и $F_{SR} \propto Sim(\Omega, \omega)$. И задача сводится к оценке $Dis(U, E)$, что на основе модели регрессии гауссовского процесса (GPR) окончательно сформулировано как

$$Dis(U, E) = GPR(F_{MI}, F_{CS}, F_{SR}) \quad (1)$$

В этой функции, мы предполагаем, что наиболее высокое сходство между источником и целевым событием по всем трём функциям - наименьшее расстояние между источником и событием, таким образом это приводит к более высокой вероятности того, что пользователь может быть на месте. На рисунке 1 представлена предлагаемая структура состоящая из четырех компонентов:

1. сбор данных и предварительная обработка;

2. моделирование социальных событий;
3. моделирование пользователя;
4. оценка расстояния между социальными событиями.



Рис. 1 – Структура системы

II. АЛГОРИТМ ПРЕДПОЛОЖЕНИЯ

Для оценки расстояния между источником сообщения о событии и событием предложена функция регрессии гауссовского процесса (FEGOR). Рабочий процесс FEGOR представлен на рисунке 2.

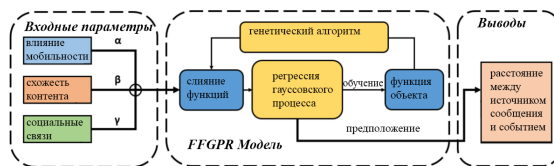


Рис. 2 – Схема процесса FAGOR

Предыдущие исследования показали, что мобильность человека может быть смоделирована на основе броуновского движения [2] с длиной шага имеющей гауссовское распределение [5]. На основе этого наблюдения мы моделируем расстояние между источником сообщения и расположением социального события как Гауссовское распределение, на среднее и дисперсию которого могут влиять различные факторы. Необходимо выявить ключевые факторы из мобильности пользователя, контента сообщения и особенности социальных отношений. Следовательно, необходимо применять регрессию гауссовского процесса (GPR) [6].

III. ЗАКЛЮЧЕНИЕ

Было предложено алгоритм на основе модели регрессии гауссовского процесса (FEGOR), которая эффективно включает три параметра: влияние мобильности, сходство контента и социальные отношения, которые извлекаются из данных источников, связанных с событиями. В процессе предположения местоположения события все абсолютные данные о местоположении

источников преобразуются в относительные целью защиты конфиденциальности расположения источников.

В сравнении с другими алгоритмами, математическая модель на основе регрессии гауссовского процесса показывает наилучший результат. Сравнение приведено в таблице 1.

Таблица 1 – Сравнение алгоритмов

Метод	Параметры			Средняя ошибка км.
	Мобильность	Контент	Соц. связи	
FEGOR	Да	Да	Да	1.7
Ryoo [3]	Да	Нет	Да	10
OLIM [4]	Да	Да	Нет	28.5

Эта работа основывается на предыдущих исследованиях на основе местоположения [7,8] и фокусируется на поиске пользователей в местах общественных мероприятий, что может иметь огромную ценность для многих социально-ориентированных приложения. Например, данные могут быть полезны правительственными чиновниками, чтобы лучше понять модели различных общественных мероприятий (например, парад, протест) и разработать лучшие стратегии управления толпой. И на основе показания обнаруженных на месте участников, правительственные или полицейские силы могут получить более подробную и значимую информация о произошедшем. Данные алгоритмы могут использоваться для любых социальных сетей, в том числе, русскоязычных.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Cheng Z, Caverlee J, Lee K. You are where you tweet: a content-based approach to geo-locating twitter users. Proceedings of the 19th ACM international conference on Information and knowledge management. ACM, 2010: 759-768.
2. Camp, Tracy, Jeff Boleng, and Vanessa Davies. "A survey of mobility models for ad hoc network research." Wireless communications and mobile computing 2.5
3. Ryoo K M, Moon S. Inferring Twitter user locations with 10 km accuracy. Proceedings of the companion publication of the 23rd international conference on World wide web companion. International World Wide Web Conferences Steering Committee, 2014: 643-648.
4. Yamaguchi Y, Amagasa T, Kitagawa H, et al. Online User Location Inference Exploiting Spatiotemporal Correlations in Social Streams. Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. ACM, 2014: 1139-1148.
5. Jiang, Bin, Junjun Yin, and Sijian Zhao. "Characterizing the human mobility pattern in a large street network." Physical Review E 80.2 (2009): 021136. (2002): 483-502.
6. Rasmussen, C. E., and Williams, C. K. I. Gaussian processes for machine learning. The MIT Press, 2006
7. Wang Y, Kankanhalli M S. Tweeting Cameras for Event Detection. Proceedings of the 24th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2015: 1231-1241.
8. Pan B, Zheng Y, Wilkie D, et al. Crowd sensing of traffic anomalies based on human mobility and social media. Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2013: 344-353.

ПРОГРАММНЫЙ ПАКЕТ RNAEXPLORER ДЛЯ ПРЕДСКАЗАНИЯ ВАРИАНТОВ АЛЬТЕРНАТИВНОГО СПЛАЙСИНГА В ПЕРВИЧНЫХ МРНК ХИМЕРНОГО ОНКОГЕНА RUNX1/RUNX1T1 ЧЕЛОВЕКА

Яцков Н. Н., Скакун В. В., Гринев В. В.

Кафедра системного анализа и компьютерного моделирования, кафедра генетики, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: {yatskou, skakun}@bsu.by

В работе представлен программный пакет RNAexploreR для предсказания событий альтернативного сплайсинга в первичных мРНК онкогена RUNX1/RUNX1T1 человека. Выполнен анализ набора экспериментальных данных для гена RUNX1/RUNX1T1 с использованием базового и распараллеленного вариантов программного пакета.

ВВЕДЕНИЕ

Гибридный онкоген RUNX1-RUNX1T1 является одним из факторов инициации и формирования острого миелоидного лейкоза у детей и взрослых. Данный онкоген кодирует более 130 изоформ РНК, закономерности образования которых до конца не установлены [1]. Современные работы нацелены на выяснение принципов, по которым идет комбинаторика экзонов во время сплайсинга и на разработку алгоритмических и программных средств для анализа и предсказания разнообразных вариантов РНК [2]. Однако интерактивные, доступные для широкого круга пользователей программные приложения, реализующие разработанные алгоритмы, практически отсутствуют или представлены локально в ограниченном виде.

Целью работы является разработка интерактивного высокопроизводительного программного пакета для предсказания вариантов альтернативного сплайсинга в первичных мРНК онкогена RUNX1/RUNX1T1 человека.

I. ПРОГРАММНЫЕ СРЕДСТВА ДЛЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

В настоящее время в открытом доступе предоставлено большое количество программных средств интеллектуального анализа данных, среди которых можно выделить: WEKA, Tanagra, Rapid Miner, KNIME, Python- и R-платформы [3]. Достоинствами того или иного программного ресурса являются: вычислительная производительность, широкий набор подключаемых библиотек, кроссплатформенность, возможность выполнения распараллеленных вычислений и работы напрямую с базами и хранилищами данных.

Основным преимуществом среды статистического программирования R является возможность использования огромного набора биоинформационных алгоритмов, алгоритмов интеллектуального анализа данных, разнообразных

статистических вычислительных ресурсов научного сообщества [4, 5]. Главным недостатком является невысокая вычислительная производительность, однако данное ограничение можно частично или полностью устранить с помощью использования процедур распараллеливания вычислений или подключения библиотек высокопроизводительных математических вычислений (например, библиотек Microsoft R Open (MRO) и Intel Math Kernel Library (MKL)). Наиболее популярные пакеты для разработки пользовательских интерфейсов программных приложений, интегрирующие R-коды, это: gWidgets, rpanel, svDialogs, RGtk2, qtbase, tcltk [5]. Новое направление в разработке R-приложений связано с созданием «реактивных» веб-интерфейсов с использованием пакета Shiny и размещением программной реализации на ресурсе shinyapps.io, предоставляемом разработчиками открытого программного обеспечения RStudio. Достоинством данного подхода является возможность удаленной работы с приложением широкой научной аудитории пользователей в режиме on-line через глобальную сеть Internet.

Для реализации программного приложения в работе выбраны вычислительная среда R и пакет Shiny для создания веб-интерфейса разработанного приложения.

II. РАСПАРАЛЛЕЛИВАНИЕ ВЫЧИСЛЕНИЙ

Научным сообществом разработано большое количество R-пакетов для реализации распараллеленных вычислений, среди которых можно выделить пакеты Rmpi, snow, snowfall, foreach и parallel [5]. Для автоматизации и упрощения работы пользователя на основе пакета snow и модели распараллеливания SOCKETS разработан пакет foreach, позволяющий относительно просто реализовать решаемую задачу в виде независимых циклов вычислений на многоядерном процессоре.

Для реализации в работе выбран пакет foreach, т.к. предоставляет: 1) логическую чет-

ко выстроенную структуру для распараллеливания кода; 2) эффективные автоматизированные средства распараллеливания; 3) возможность распараллеливания на многоядерном процессоре; 4) возможность анализа больших массивов данных, реализуемую с помощью пакета *iterators*; 5) набор примеров по распараллеливанию задач интеллектуального анализа данных.

III. ОПИСАНИЕ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

Тестирование пакета произведено на двух вычислительных системах (персональных компьютерах), далее представленных как ПК1 и ПК2, характеризующихся различными вычислительными возможностями и производительностью. ПК1 содержит двухядерный процессор DualCore Intel Pentium E5700, работает под 32 битной OS MS Windows XP. Данный ПК используется как эталонный вариант компьютера. ПК2 является более продвинутой вычислительной системой в смысле распараллеливания решаемых задач, имеет четырехядерный процессор QuadCore Intel Core 2 Quad Q9400, работает под 64 битной OS MS Windows 10. ПК2 является хорошим примером компьютерного (4-х ядерного) кластера для решения задач распараллеливания.

В работе исследуются способы повышения эффективности пакета с помощью подключения высокопроизводительной библиотеки Intel MKL и распараллеливания программного кода.

IV. РЕЗУЛЬТАТЫ

Разработан программный пакет RNAexploreR. On-line версия пакета доступна по ссылке https://dsa-sm.shinyapps.io/NIR_bio_code_Sh/. Главное окно интерфейса пакета состоит из набора панелей, соответствующих основным этапам анализа данных. На каждом этапе анализа пользователь должен загрузить требуемый файл данных (файлы экзонов, экспериментально подтвержденных транскриптов, теоретически предсказанных транскриптов) и установить системные параметры алгоритмов интеллектуального анализа данных. Пример результирующего этапа анализа данных представлен на рис. 1.

Выполнено исследование эффективности работы базового и распараллеленного вариантов программного пакета на примере анализа полного набора экспериментальных данных для гена RUNX1/RUNX1T1 [1]. Проводилось от 3 до 6 запусков пакета на каждом из ПК.

В результате распараллеливания алгоритмов удалось существенно сократить время работы пакета на 41% на ПК1 (с 8.66 до 5.13 мин) и на 316% ПК2 (с 6.84 до 1.80 мин). Распараллеленный вариант пакета работает в 5 раз быстрее на ПК2, чем его нераспараллеленная Shiny-версия на ПК1.

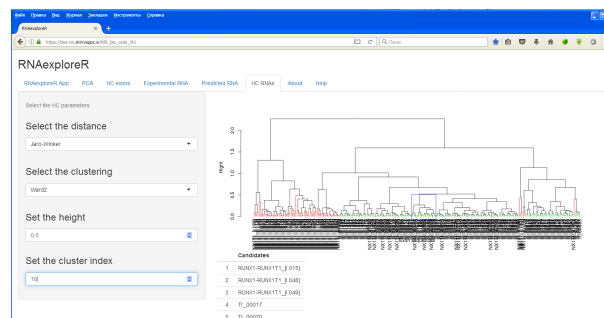


Рис. 1 – Результаты работы алгоритма иерархической кластеризации пула уникальных экспериментальных (красный цвет) и теоретических (зеленый цвет) транскриптов онкогена RUNX1/RUNX1T1. Для демонстрации результата список теоретически предсказанных транскриптов сокращен до 3000

V. ЗАКЛЮЧЕНИЕ

Проведено исследование существующих свободных программных средств интеллектуального анализа данных для реализации программного приложения для предсказания событий альтернативного сплайсинга в первичных мРНК химерного онкогена RUNX1/RUNX1T1 человека. Выбраны наиболее оптимальные программные средства для реализации исследуемой задачи – вычислительная среда R и пакет Shiny для создания веб-интерфейса приложения. Разработано веб-приложение и распараллеленный вариант пакета RNAexploreR. Выполнен анализ набора данных для гена RUNX1/RUNX1T1 с использованием базового и распараллеленного вариантов программного пакета на различных компьютерах.

VI. СПИСОК ЛИТЕРАТУРЫ

1. Ильюшёнко, И. Н. Изучение закономерностей сплайсинга РНК гибридного онкогена RUNX1-RUNX1T1 человека с помощью методов интеллектуального анализа данных и высокопроизводительного секвенирования / И. Н. Ильюшёнко, Е. П. Гулько, М. Л. Антонович, Н. Н. Яцков, А. М. Кустанович, А. Ю. Сухаревский, В. В. Скакун, В. В. Гринев // Молекулярная и прикладная генетика. – 2017. – Том 23. – С. 92–101.
2. Grinev, V. V. Decoding of exon splicing patterns in the human RUNX1-RUNX1T1 fusion gene / V. V. Grinev, A. A. Migas, A. D. Kirsanova, O. A. Mishkova, N. Siomava, T. V. Ramanouskaya, A. V. Vaitsiankova, I. M. Ilyushonak, P. V. Nazarov, L. Vallar, O. V. Aleinikova // Int. J. Biochem. Cell Biol. – 2015. – Vol. 68. – P. 48–58.
3. Kdnuggets [Electronic recourse] / – Mode of access: <https://www.kdnuggets.com>. – Date of access: 25.08.2018.
4. Bioconductor [Electronic recourse] / – Mode of access: <http://www.bioconductor.org>. – Date of access: 25.08.2018.
5. The R Project for Statistical Computing [Electronic recourse] / – Mode of access: <http://www.r-project.org>. – Date of access: 25.08.2018.

СТАТИСТИЧЕСКИЙ ПОДХОД К ПРЕДСКАЗАНИЮ СОБЫТИЙ АЛЬТЕРНАТИВНОГО СПЛАЙСИНГА В ПЕРВИЧНЫХ МРНК ГЕНОВ ЧЕЛОВЕКА

Яцков Н. Н., Скакун В. В., Гринев В. В.

Кафедра системного анализа и компьютерного моделирования, кафедра генетики, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: {yatskou, skakun}@bsu.by

Разработан статистический подход к предсказанию событий альтернативного сплайсинга в первичных мРНК генов человека. На основе исследованных алгоритмов снижения размерности данных, иерархического кластерного анализа, вычисления расстояний между символьными последовательностями выполнено сравнение экспериментально подтвержденных транскриптов модельных генов человека. Точность предсказания разработанного подхода составляет 90-95% для рассмотренных пар модельных генов.

ВВЕДЕНИЕ

Конститутивный и альтернативный сплайсинг является фундаментальным процессом, протекающим во всех без исключения клетках эукариот и приводящим к образованию зрелых функциональных РНК-продуктов [1]. Однако, принципы (правила) комбинаторики экзонов во время сплайсинга до сих пор не установлены [2]. Следует отметить об ограниченном применении или даже полном отсутствии стандартов или единых систематизированных статистических подходов к анализу и интерпретации возможных экзонных последовательностей генов человека. Для выяснения принципов, по которым идет комбинаторика экзонов во время сплайсинга [3], требуется разработка статистических алгоритмов и программных средств для анализа и предсказания разнообразных вариантов генерации РНК.

Целью работы является разработка статистического системного подхода к анализу и предсказанию событий альтернативного сплайсинга в первичных мРНК генов человека. Реализованы и исследованы наиболее эффективные алгоритмы интеллектуального анализа данных. Проверка работоспособности разработанных алгоритмов выполнена на примере набора модельных генов человека [4].

I. ОБЪЕКТ И ПРЕДМЕТ ИССЛЕДОВАНИЯ

Объект исследования – альтернативный сплайсинг РНК онкогена человека. В качестве примера рассмотрен набор из 14 произвольно отобранных негомологичных генов человека [4]. Для каждого из рассмотренных генов из базы данных Ensembl получены списки уникальных экзонов и экспериментально подтвержденных транскриптов. Предметом исследования являются алгоритмы интеллектуального анализа данных, позволяющие предсказать события альтернативного сплайсинга в первичных РНК генов человека.

II. МЕТОДИКА СТАТИСТИЧЕСКОГО ПОДХОДА К ПРЕДСКАЗАНИЮ СОБЫТИЙ АЛЬТЕРНАТИВНОГО СПЛАЙСИНГА

Идея статистического подхода состоит в снижении размерности пространства экзонных признаков и объединении близко расположенных экзонов в ограниченное число классов, замене экзонных путей генерации РНК на последовательности соответствующих меток классов экзонов, вычислении расстояний между транскриптами РНК, объединении близкорасположенных объектов РНК в сходные кластеры. Основные этапы методики подхода с учетом выбранных наиболее оптимальных алгоритмов интеллектуального анализа данных:

Этап 1. Анализ полного набора признаков экзонов с использованием метода главных компонент [5]. Шкалирование и центрирование данных. Отбор главных компонент, объясняющих 95% вариации в данных.

Этап 2. Иерархическая кластеризация [5] экзонов гена на основе отобранного набора новых признаков (главных компонент). Разбиение экзонов на кластеры и сопоставление каждому кластеру уникального индекса в символах латинского алфавита (от а до z).

Этап 3. Преобразование символов последовательностей транскриптов РНК (от имен экзонов) к меткам кластеров, в которых расположены соответственные экзоны.

Этап 4. Удаление транскриптов дубликатов.

Этап 5. Вычисление расстояний между транскриптами РНК. Иерархическая кластеризация пула уникальных (в смысле не дубликатов) транскриптов. Представление результатов анализа в виде дендрограммы.

III. ОПИСАНИЕ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

Для проверки работоспособности разработанных алгоритмов подхода рассмотрено предсказание событий сплайсинга на примерах экс-

периментально подтвержденных транскриптов различных пар модельных генов. В случае успешной работы подхода транскрипты РНК различных генов должны быть предсказаны с высокой точностью.

Для определения наиболее оптимального выбора метрического расстояния и метода связывания объектов в алгоритмах иерархического кластерного анализа рассмотрен кофенетический корреляционный коэффициент [5]. В качестве мер сравнения последовательностей транскриптов РНК генов исследованы различные алгоритмы лексиграфического анализа, такие как расстояния Левенштейна, Дамерау-Левенштейна, наибольшей подстроки, q-грамм, Джакарда, Джаро и Джаро-Винклера [6]. Для оценки точности предсказания принадлежности транскриптов РНК к заданному гену используется точность классификации (в %) $A = 100 \cdot (N1 + N2) / N$, где N1 и N2 — число правильно классифицированных транскриптов для двух генов, N — общее число транскриптов двух генов.

IV. РЕЗУЛЬТАТЫ

В ходе анализа полного набора признаков экзонов с использованием метода главных компонент отобрано менее 100 наиболее значимых компонент (из более чем 1400 исходных признаков экзонов), объясняющих 95% вариации в данных. Иерархический кластерный анализ с использованием расстояния Минковского и связывания кластеров по Варду [5] является наиболее эффективным как для группировки экзонов генов, так и для нахождения классов схожих транскриптов.

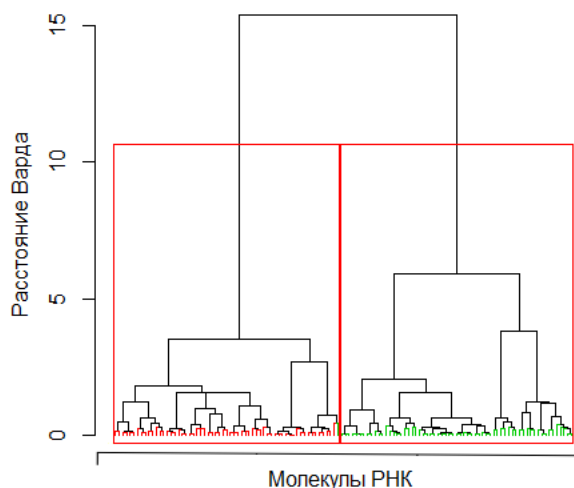


Рис. 1 – Результаты работы алгоритма иерархической кластеризации пула уникальных экспериментальных транскриптов двух модельных генов (красный и зеленый цвет)

Расстояние Джаро-Винклера с оптимизацией параметра штрафа является наилучшим, обеспечивая более высокую точность кластеризации при меньшем числе используемых кластеров. Точность разделения составляет 90-95% для рассмотренных пар модельных генов.

Таким образом, на примерах модельных генов подтверждена работоспособность разработанного подхода: экспериментально подтвержденные транскрипты пар модельных генов разделяются на два класса (рис. 1).

V. ЗАКЛЮЧЕНИЕ

Впервые предложен и исследован на модельных генах человека статистический подход к сравнению транскриптов генов человека, основанный на применении алгоритмов снижения размерности данных, иерархического кластерного анализа, сравнения символьных последовательностей. Точность предсказанию событий альтернативного сплайсинга составляет более 90% для рассмотренных пар модельных генов. Предложенные алгоритмы могут быть использованы для изучения организации и функционирования как aberrантных, так и нормальных генов человека, а получаемые при этом данные могут быть полезны для дифференциальной диагностики и построения прогноза течения заболеваний, имеющих генетическую природу.

VI. СПИСОК ЛИТЕРАТУРЫ

1. Hang, J. Structural basis of pre-mRNA splicing / J. Hang, R. Wan, C. Yan, Y. Shi // *Science* – 2015. – Vol. 349, № 6253. – P. 1191–1198.
2. Barash, Y. Deciphering the splicing code / Y. Barash, J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe, B. J. Frey // *Nature* – 2010. – Vol. 465, № 729. – P. 53–59.
3. Grinev, V. V. Decoding of exon splicing patterns in the human RUNX1-RUNX1T1 fusion gene / V. V. Grinev, A. A. Migas, A. D. Kirsanova, O. A. Mishkova, N. Siomava, T. V. Ramanouskaya, A. V. Vaitiankova, I. M. Ilyushonak, P. V. Nazarov, L. Vallar, O. V. Aleinikova // *Int. J. Biochem. Cell Biol.* – 2015. – Vol. 68. – P. 48–58.
4. Aken, B. L. The Ensembl gene annotation system / B. L. Aken, S. Ayling, D. Barrell, L. Clarke, V. Curwen, S. Fairley et al. // *Database (Oxford)* – 2016. – Database URL: <http://www.ensembl.org/index.html>.
5. Интеллектуальный анализ данных / Н. Н. Яцков – Минск: БГУ, 2014. – 151 с.
6. Cohen, W. A comparison of string metrics for matching names and records / W. W. Cohen, P. Ravikumar, S. E. Fienberg // *KDD Workshop on Data Cleaning and Object Consolidation* – 2003. – Vol.3. – P. 73–78.

ПРИМЕНЕНИЕ МОДИФИЦИРОВАННЫХ АЛГОРИТМОВ JAVA ANPR ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ НОМЕРОВ АВТОМОБИЛЕЙ

Заерко Д. В., Липницкий В. А.

Кафедра информатики, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: zaerko1991@gmail.com, valipnitski@yandex.by

В работе описаны основные этапы автоматического распознавания номеров автомобилей, а также новшества, задействованные для каждого из этапов; кратко представлены условия работы алгоритмов и основные предпосылки для изменения базовых алгоритмов. Предметом исследования выступают модифицированные алгоритмы определения области выделения, нормализации и распознавания номера. Основной целью работы является обоснование выбора наилучшего пакета программного обеспечения, позволяющего успешно провести распознавание номеров, а так же, описание некоторых результатов полученных после внесения изменений в базовые алгоритмы пакета JavaANPR.

I. АВТОМАТИЗАЦИЯ КОНТРОЛЯ ЗА ТРАНСПОРТНЫМИ ПОТОКАМИ

Последнее столетие мы наблюдаем интенсивное развитие автомобильной инфраструктуры и экспоненциальный рост числа участников движения на дорогах. Это неизбежно приводит к росту количества ДТП, правонарушений и преступлений на дорогах. Невозможность оперативной обработки информации, используя сугубо человеческие ресурсы, снижает как по объему, так и по качеству контроль над потоками транспортных средств. Как и во всех сферах человеческой деятельности, здесь необходимо применение современных средств и систем IT-технологий.

Одним из наиболее распространённых средств контроля над автомобильным транспортом является индикация через уникальные номерные знаки, предоставляемые для каждого автомобиля и закрепленные за определенными лицами или организациями в государственном реестре. Автоматическая индикация номеров должна проводиться посредством оптического устройства, не зависимо от искажающих факторов, таких как: физические особенности, погодные условия, применение отражающих средств и т.д., а так же с минимальным участием человека. Минимизация использования человеческого труда и автоматизация приведёт к сокращению расходов автотранспортных предприятий и предприятий, сфера деятельности которых связана с охраной и безопасностью. Для решения этой проблемы предполагается:

1. внедрение новых автоматизированных устройств слежения за транспортными потоками (качеством съемки),
2. повышения скорости передачи и объема полученных визуальных данных (пропускная способность узлов и каналов связи),
3. использования новых алгоритмов очистки, распознавания и классификации полученных данных.

Если качество съемки и пропускная способность зависит в большей степени от физических свойств устройства, а их качество все время повышается, то алгоритмы зачастую заметно отстают от требований времени. Это связано с математической сложностью их создания, а также с отставанием их последующей проверки их качества на практике.

II. ВЫБОР БИБЛИОТЕКИ И ОСНОВНЫЕ ЭТАПЫ РАСПОЗНАВАНИЯ

Задача распознавания автомобильных номеров не нова и уже имеет множество систем, реализующих её. Существуют как частные реализации так и узкоспециализированные. Большинство частных реализаций, основано на комбинировании двух популярных библиотек: OpenCV (Open Source Computer Vision Library) свободно распространяемой под лицензией BSD и TesseractOCR (Tesseract engine for optical character recognition), так же свободно распространяемая но, под лицензией Apache 2.0 и поддерживаемая корпорациями Google и Hewlett-Packard. Обе библиотеки включают алгоритмы, предполагающие их машинное обучение (Machine Learning) в процессе работы на обучающих выборках.

Пакет OpenCV, обширнейший набор разнообразных алгоритмов, как правило, используются для предварительной обработки изображения, например, локализации номера на изображении, а также для очистки изображения от помех. TesseractOCR используется для непосредственного распознавания номера в ограниченной области. Хотя эти библиотеки и являются универсальными, но имеют свои преимущества для решения специфических задач. Так, например, алгоритмы TesseractOCR очень чувствительны к помехам, искажающим изображения, однако позволяют качественнее провести их обучение на относительно небольшой обучающей выборке.

Есть также узкоспециализированные библиотеки и фреймворки на различных языках программирования, непосредственно предназначенные для распознавания автомобильных номеров. Значительный интерес представляет система ANPR (Automatic number-plate recognition), алгоритмы которой реализованы на наиболее популярных языках программирования, таких как: C# (проект Opos), Java (проект JavaANPR) и т.д.

Рассмотрим реализацию проекта JavaANPR [1]. Основное преимущество этой библиотеки в ее кроссплатформенности. Кроме этого, все алгоритмы написаны на Java без использования нативных библиотек, что сильно упрощает их применение. Так же эту библиотеку с определенной доработкой можно использовать на устройствах под управлением OS Android. Отметим, что это отдельный проект, который использует свои уникальные алгоритмы. При выборе инструмента для распознавания, немаловажную роль сыграло так же то, что библиотека является свободной для распространения и модификации текстов программ.

Разрабатываемая нами система автоматического распознавания номеров автомобилей (далее CAPNA) должна содержать решение следующих задач:

- предварительный поиск номера – обнаружение области, в которой содержится номер;
- нормализация номера – определение точных границ номера, нормализация контраста;
- распознавание текста – чтение всего, что нашлось в нормализованном изображении.

Система JavaANPR имеет свои уникальные реализации для решения каждой из задач, однако позволяет включать в работу системы собственные или видоизмененные алгоритмы для каждого из пунктов, что и было сделано. Основной практической целью была апробация видоизмененных алгоритмов и тестового модуля контроля и обучение системы в следующих простейших, но являющихся базовыми, случаях:

1. алгоритмы поиска номера применяются с предположением о том, что набор символов (знаки автомобильного номера) присутствуют в единственной области изображения;

2. границы изображения – не «размыты» и нет значительного наклона относительно центра изображения;
3. для распознавания символов текста используются незначительно зашумленные изображения, не предполагающие применение особых фильтров для их очистки.

III. МОДИФИКАЦИЯ АЛГОРИТМОВ И РАЗВИТИЕ СИСТЕМЫ

Обучение проводилось на номерах с латинскими символами, а объем выборки для обучения составил относительно небольшую коллекцию из 2500 изображений. Кроме того, первоначальная система была дополнена специальным модулем, проводящим контроль распознанного изображения в соответствии с набором условий: ограниченность числа символов в номере, наличие только латинских символов и только арабских цифр. В случае возникновения ошибок цикл распознавания номера повторялся до тех пор, пока либо не были соблюдены все условия, либо число итераций не достигало критического значения (например, максимально 10 циклов). Опытным путем было обнаружено, что для хорошо обученного алгоритма, как правило, хватало 2-3 циклов распознавания. Модифицированный алгоритм поиска объекта на изображении учитывал на каждом шаге «успешность» распознавания наперед идущего, с целью пошагового подбора наилучшего набора точек.

В дальнейшем изучение и применение модифицированных алгоритмов системы будет продолжено на изображениях с помехами. Например: изменение угла и степени освещения, наличие элементов, затрудняющих распознавание (грязь, отражение от капель воды), а так же удаленности объекта на изображении.

Кроме того, планируется увеличить набор условий для контрольного модуля, применяемого на последнем этапе распознавания, а также провести сравнения модифицированных и базовых алгоритмов на скорость распознавания. В качестве дополнения предполагается, что система будет проводить анализ распознанной информации по некоторым критериям: регион, спец номер и т. д.

1. Martinsky O. Algorithmic And Mathematical Principles Of Automatic Number Plate Recognition Systems / O. Martinsky. // B.Sc. thesis, Department of Intelligent Systems, Faculty of Information Technology, Brno University of Technology. – 2007. – P. 1.

ФИЛЬТРАЦИЯ СПАМ-ПИСЕМ С ПОМОЩЬЮ АЛГОРИТМОВ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Заграй В. Ю., Гуринович А. Б.

Кафедра информационных технологий автоматизированных систем, кафедра вычислительных методов и программирования, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: v_vld@mail.ru, gurinovich@bsuir.by

Процесс отсеивания спама является неотъемлемой частью информационной безопасности. Он достаточно ресурсоемок и нетривиален. Для оптимизации процесса защиты информации используются нейронные сети, которые становятся частью существующей системы защиты. Это одно из самых перспективных направлений в области защиты данных.

ВВЕДЕНИЕ

Одной из главных задач при осуществлении обмена сообщениями и письмами является фильтрация или отсеивание вредоносных сообщений и так сообщений, не содержащих информации, не требующей внимания пользователя, как реклама ненужных товаров или предложения для нелегального заработка, называемые спам-письмами. Алгоритмы фильтрации спам-писем предназначены для отсеивания таких сообщений, но имеют вероятность ложного срабатывания.

Цель работы состоит в нахождении способа уменьшения вероятности пропуска спам-писем, отправляемых пользователям посредством сети интернет. Последующее применение результатов работы позволит уменьшить указанную ранее вероятность.

I. ПОДГОТОВКА ДАННЫХ

Для использования алгоритмов фильтрации для определения спам-писем, необходимо определить ключевые характеристики писем, по которым производится фильтрация.

Таковыми характеристиками являются:

- Повторяющиеся слова;
- словосочетания;
- общие количества символов, знаков пунктуации, цифр, пробельных символов, слов;
- частота каждой буквы, специальных символов;
- средние длины слова, предложения;
- распределение частоты длин слов;
- мера разнообразия;
- количество уникальных слов.

Для получения характеристик используется словарь, содержащий в себе слова, символы и прочие элементы текста, наиболее характерные для спама.

Выбор ключевых характеристик напрямую влияет на выбор конкретного алгоритма фильтрации. У некоторых алгоритмов конечный результат зависит от начальных данных, т.е. при неправильном выборе характеристик писем можно увеличить вероятность ложного срабатывания. Исходя из этих характеристик, можно оце-

нивать текст на принадлежность к спаму комплексно, полагаясь на множество разнородных параметров, которые дополняют друг друга и уточняют оценку при принятии решения.

II. ВЫБОР АЛГОРИТМОВ

Для выполнения фильтрации спам-писем необходимо полученные ранее характеристики сообщений обработать алгоритмами классификации текста и данных. Существуют классические алгоритмы и алгоритмы на основе искусственных нейронных сетей.

Классические алгоритмы представляют собой методы и алгоритмы построенные на применении методов статистического анализа данных и математических вычислений. К ним относятся:

- Наивный байесовский классификатор;
- метод k-ближайших соседей;
- метод опорных векторов;
- генетические алгоритмы.

Алгоритмы фильтрации на основе искусственных нейронных сетей предназначены на разделения задачи на блоки. При этом каждый блок обрабатывается более простыми процедурами, чем в классических алгоритмах, что позволяет динамически расширять сети без нарушения последовательности выполнения алгоритма.

К этим алгоритмам относятся:

- Распознавание образов;
- перцептрон;
- нейронная сеть Кохонена;
- самоорганизующаяся карта Кохонена.

Применение алгоритмов на основе необученных нейронных сетей приводит к нежелательным результатам, т.е. к пропуску спам-писем. Данный недостаток возможно минимизировать путём совместного использования классических алгоритмов и алгоритмов на основе нейронных сетей на начальном этапе обучения или самообучения сети.

Как было указано в предыдущем разделе, из исходных сообщений, вероятно, содержащих спам-письма, извлекаются отдельные данные, представляющие собой количественные характеристики сообщений. Данные характери-

ки возможно обрабатывать параллельно с целью уменьшения времени выполнения фильтрации.

III. ОБРАБОТКА КЛЮЧЕВЫХ ХАРАКТЕРИСТИК

При фильтрации спам-писем применяются классические алгоритмы, так как они уже достаточно хорошо изучены. Одним из них является Наивный байесовский классификатор. Его суть состоит в определении вероятности того, что конкретное письмо является спамом. Одним из недостатков является тот факт, что необходимо заранее определить набор ключевых характеристик, описывающих спам-письма, которого лишены самообучаемые нейронные сети. Преимуществом является быстрое время выполнения. При обучении алгоритма вычисляются вероятности появления отдельных характеристик в спам-письмах формула: $P(C) = \frac{m}{n}$, где m - это количество спам-писем, n - количество всех писем с характеристикой C

Также применяются неклассические алгоритмы. Анализ ключевых характеристик нейронной сетью напоминает байесовскую фильтрацию спама, где для каждого слова или словосочетания можно установить коэффициент определения письма как спам. Однако, в отличие от байесовского фильтра, здесь коэффициенты - это веса между нейронами сети, способные динамически изменяться в процессе обучения, что позволяет эффективно обнаруживать новый и ранее неизвестный спам за счет умения нейронной сети обобщать накопленный опыт. Нейронные сети внешне похожи на Наивный байесовский классификатор, но структурно различаются.

Такую нейронную сеть можно структурно реализовать в виде многослойного перцептрона со скрытыми слоями или в виде сети Кохонена.

Перцептронная сеть проста в реализации и представляет собой перцептрон с числом входных параметров n , равных размерности входного вектора характеристик (при использовании полученных выше характеристик $n = 7$). Она имеет единственный нейрон, выдающий значение вероятности обнаружения спама в тексте, принимающий значение от 0 до 1. Перцептронная сеть выполняет единственную функцию - принятие решения о наличии спама.

Также может быть применена сеть Кохонена, которая выполняет кластеризацию данных, что позволяет эффективнее определить направленность текста, в том числе отбросить на этапе кластеризации текст, являющийся обычным текстом. Нейрона сети Кохонена в базовом варианте имеет вид:

$$y_j = w_{j0} + \sum_{i=1}^m w_{ji}x_i$$

В данной формуле w_{ji} - это вес i -й характеристики j -ого нейрона.

Обычно применяется немаленькое число нейронов для нахождения "лучшего" нейрона, имеющего наибольший выход. Полученное на выходе значение сверяется с пороговой величиной, определяющей границу между спамом и обычной почтой.

Нейронные сети, как и биологические организмы в начале своего существования, являются необученными, то есть не могут использоваться без обучения. Для минимизации ошибок от использования недостаточно обученных нейронных сетей следует применять классические алгоритмы, так как они показывают меньшую вероятность ложного срабатывания по сравнению с нейронной сетью. Однако впоследствии нейронные сети, обученные на реальных данных и результатах, полученных от классических алгоритмов, имеют меньшую вероятность ложного срабатывания при высокой производительности, чем классические алгоритмы.

ЗАКЛЮЧЕНИЕ

Среди алгоритмов, применяемых для фильтрации данных, а в частности электронной почты и сообщений, имеется множество как производительных, но имеющих высокую вероятность ложного срабатывания, так и точных, но медленных алгоритмов. Совместное использование нейронных сетей с классическими алгоритмами позволяет уменьшить количество спам-писем, а также уменьшить вероятность их пропуска фильтром.

1. Rosenblatt, F. F. Principles of Neurodynamic: Perceptrons and the Theory of Brain Mechanisms / F. F. Rosenblatt - 1965. - 480 с.
2. Spam detection using neural networks [Электронный ресурс] / Портал GitHub. -- Режим доступа: <https://medium.com/emergent-future/spam-detection-using-neural-networks-in-python-9b2b2a062272>. -- Дата доступа: 20.09.2018
3. Statistical Comparisons of Classifiers over Multiple Data Sets [Электронный ресурс] -- Режим доступа: <http://sci2s.ugr.es/sicidm/pdf/2006-Demsar-JMLR.pdf>. -- Дата доступа: 23.09.2018
4. Нейронные сети: полный курс / С. В. Хайкин [и др.]. - М: Вильямс, 2006. - 1104 с.
5. Ясницкий, Л. Н. Введение в искусственный интеллект / Л. Н. Ясницкий -М: Издательский центр «Академия», 2010. - С. 176.
6. Андреев, А. М. Автоматическая классификация текстовых документов с использованием нейросетевых алгоритмов и семантического анализа. Режим доступа: <http://www.inteltec.ru> (дата обращения: 10.09.2018).
7. Garshin, A. A. An automated system of recognizing handwritten digits based on convolution neural networks / A. A. Garshin, O. P. Soldatova - Register of computer programs - 2010
8. Головкин, В. А. Нейронные сети: обучение, организация и применение / В. А. Головкин // М.: ИПРЖР, 2001. - 205 с.

ИСПОЛЬЗОВАНИЕ ЯРКОСТНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ В ИНТЕРЕСАХ РАСПОЗНАВАНИЯ ВОЗДУШНЫХ ОБЪЕКТОВ В ОПТИКО-ЛОКАЦИОННЫХ СИСТЕМАХ

Зайко Е. В., Ярмолик С. Н., Храменков А. С., Свинарский М. В., Леонович А.С.
Кафедра автоматики, радиолокации и приемо-передающих устройств, Военная академия Республики
Беларусь
Минск, Республика Беларусь
E-mail: zayko.eugene@mail.ru, yarmsergei@yandex.ru, xras.tech@mail.ru, mechislav1993@gmail.com

В докладе рассмотрен вариант классификации воздушных объектов в оптико-локационных системах с использованием яркостных признаков изображений. Для выбранного алфавита классов исследована эффективность использования различных вариантов яркостных признаков изображений. Использование дифференциальной контрастности изображений объектов анализируемых классов позволило получить и проанализировать характеристики распознавания.

ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Развитие вычислительной техники и цифровых технологий обработки сигналов позволяет повышать качество, скорость и объемы обработки изображений в оптико-локационных системах [1]. Выделяемые в результате обработки изображения объектов обладают высокой информативностью, что открывает широкие возможности практического использования информации, получаемой в результате обработки и преобразования изображений в устройствах классификации объектов. Задача классификации (распознавания) получила широкое распространение как в гражданской, так и в военной областях применения. Вариантами использования результатов распознавания являются целераспределение, целеуказание, селекция целей, медицинская и техническая диагностика и др. [4]. Развитие локационных систем оптического диапазона, как элементов общего комплекса ПВО, неразрывно связана с определением класса наблюдаемого объекта. Несмотря на большое количество публикаций, связанных с решением задач классификации образов в оптическом диапазоне волн, вопросы описания и выделения классификационных признаков, а также разработки эффективных алгоритмов принятия решений при классификации изображений остаются весьма актуальной темой исследований.

ОСНОВНАЯ ЧАСТЬ

Математическое описание обрабатываемых изображений предполагает использование двух основных подходов: детерминированного и статистического [1]. При детерминированном описании изображения используется математическая функция, характеризующая анализируемые свойства изображения в каждой точке. При статистическом описании изображение определяется своими усредненными статистическими характеристиками [5].

В наиболее общем случае изображение описывается с помощью функции распределения яркости ($F(x, y, t)$), или другой физической величины, связанной с яркостью [1]. Часто анализ изображения производят в частотной области, используя двумерное преобразование Фурье. В ряде практических случаев изображение описывается совместной плотностью вероятности значений функции $F_j(x_j, y_j)$ в точках отсчета x_j, y_j . В качестве моделей плотности вероятности яркости применяются различные плотности распределения вероятностей (Гауссовская, Релея, экспоненциальное и др.) [2]. Изображение в цифровой форме представляет собой матрицу, размерностью $M \times N$, где каждый элемент матрицы характеризует значение яркости изображения. Каждый компонент изображения характеризуется текущим значением яркости. Случайные значения яркости изображения объекта и фона целесообразно характеризовать средним значением яркости (МО) и величиной отклонения от среднего уровня (СКО). В интересах сопоставительного анализа компонентов изображения различной яркости удобно использовать значение дифференциальной контрастности: ΔK .

Классификационным признаком изображения объекта называют его определенное свойство, отражающее характерные отличия изображения объекта [1]. В работе рассмотрен один из вариантов классификации признаков объектов, наблюдаемых оптико-локационными системами. К классификационным признакам относят, например, яркостные различия в компонентах изображения или геометрические различия объектов на изображениях. Для формирования отличительных признаков могут использоваться типовые математические преобразования (преобразование Фурье, Мелина, интегральное описание изображения и т.п.), либо интуитивный подход (формирование масок, нейросетевые методы выделения признаков) [1, 3].

Анализ эффективности используемых отличительных признаков изображений оптических объектов предполагает выбор показателей, характеризующих качество решения задачи классификации. Задача распознавания изображений объектов, наблюдаемых оптико-локационными системами, является статистической. Качество решения любой статистической задачи необходимо оценивать с помощью вероятностных методов [4]. В качестве показателей качества, характеризующих эффективность принимаемых решений о классе изображения объекта, в работе использованы вероятности правильных D_k и ложных F_k решений о классе объекта. В качестве аналога отношения сигнал-шум использован коэффициент дифференциальной контрастности ΔK характеризующий степень различия в значениях уровня яркости фона и изображения объекта. Для получения выбранных показателей качества использовался метод математического моделирования. Выбор метода обусловлен простотой его применения, а также удобством формирования статистики, необходимой для оценивания характеристик.

Анализ эффективности выбранных классификационных признаков осуществлялся с помощью разработанного моделирующего комплекса. Данный комплекс обеспечивает выделение требуемых классификационных признаков изображения наблюдаемого объекта, обеспечивает расчет требуемой метрики, а также производит оценку эффективности принимаемых решений. Моделирующий комплекс, реализованный в среде программирования C++ Builder имеет гибкую и открытую архитектуру, обеспечивающую возможность модернизации и совершенствования его составных частей. Представлены результаты исследования аэродинамических объектов 3 типовых классов: самолет, крылатая ракета, вертолет. В интересах анализа использовались следующие отличительные признаки изображений выбранных классов: - гистограммные признаки в совокупности с геометрической формой изображения объекта;- интегральные признаки с учетом значений яркости изображения объекта.

Применительно к заданным условиям наблюдения объектов формировались случайные реализации анализируемых изображений. В устройстве распознавания воспроизводились адаптивные эталонные изображения. В интересах определения класса изображения наблюдаемого объекта рассчитывались коэффициенты правдоподобия анализируемых гипотез. Решение о классе наблюдаемого объекта принималось на основе оцениваемых функций расстояния между анализируемыми векторами признаков (расстояние Кульбака-Лейблера). Решение принималось в пользу наиболее правдоподобной гипотезы. После подсчета количества принятых реше-

ний о классе распознаваемого объекта рассчитывались вероятностные показатели качества (D и F)

Результаты моделирования, представлены на рисунке в виде графических зависимостей вероятностей правильного и ложного распознавания от значения дифференциальной контрастности для анализируемых яркостных признаков.

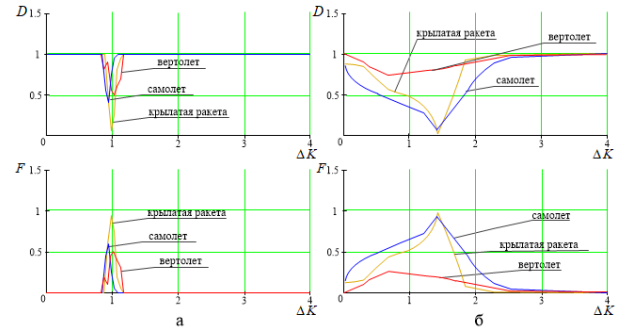


Рис. 1 – Показатели качества распознавания: а – для гистограммных признаков; б – для интегральных признаков

ЗАКЛЮЧЕНИЕ

Полученные графики позволяют утверждать, что при использовании гистограммных признаков полученные значения вероятностных показателей качества имеют более высокие значения, чем при использовании геометрических признаков. Следует также отметить, что приведенные результаты исследований не позволяют в полной мере характеризовать информативность анализируемых классификационных признаков. Многогранность задачи распознавания изображений предполагает комплексный учет большого количества сопутствующих факторов, влияющих на процесс принятия результирующего решения. Вместе с этим, важно подчеркнуть, что разработанный и программно реализованный комплекс моделирования позволяет решать ряд задач, связанных с оцениванием эффективности классификационных признаков в различных условиях наблюдения.

1. Анисимов, Б. В. Распознавание и цифровая обработка изображений: учеб. пособие / Б. В. Анисимов, В. Д. Курганов, В. К. Злобин, — М.: Высш. шк., 1983. —295 с.
2. Прэйт, У. Цифровая обработка изображений / У. Э. Прэйт. Пер. с англ. —М.: Мир, 1982. —Кн.1 —312 с.
3. Прэйт, У. Цифровая обработка изображений / У. Э. Прэйт. Пер. с англ. —М.: Мир, 1982. —Кн.2 —380 с.
4. Охрименко, А. Е. Основы извлечения, обработки и передачи информации. Ч.3. Распознавание – различие сигналов / А. Е. Охрименко. — Минск.: ВГУИР, 1994. —64 с.
5. Лабунец, Л. В. Цифровые модели изображений целей и реализаций сигналов в оптических локационных системах. Учеб. пособие / Л. В. Лабунец. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2007. —216 с.

ОПРЕДЕЛЕНИЕ ТРЕНДОВОЙ МОДЕЛИ ТИПА СДВИГА ДЛЯ НЕСТАЦИОНАРНОГО ВРЕМЕННОГО РЯДА

Жук Е. Е., Чернявский Д. В.

Кафедра математического моделирования и анализа данных, Белорусский государственный университет
Минск, Республика Беларусь

E-mail: zhukee@mail.com, lobiback@gmail.com

Данная работа посвящена проблеме определения ближайшей к реализации нестационарного временного ряда трендовой модели. Причем здесь трендовые модели задаются каждая своим типовым трендом, образуя семейства сдвига.

ВВЕДЕНИЕ

Во многих приложениях, в том числе прогнозирование поведения цены на бирже [4], важно спрогнозировать не само значение цены, а ее тенденцию, которая определяется базовыми трендами. Здесь исследуется проблема определения трендовой модели, ближайшей к наблюдаемой реализации (например, цены). Для принятия решений предлагается использовать решающее правило по методу наименьших квадратов. Для него вычислен риск, а также в качестве примера рассмотрен случай двух альтернативных трендовых моделей.

I. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И ПОСТАНОВКА ЗАДАЧИ

Пусть имеется реализация $X = \{x_t\}_{t=1}^T$ длительности T нестационарного временного ряда (ВР). Отсчеты $x_t \in R$, $t = \overline{1, T}$, этого ВР расположены возле своего тренда [2]:

$$x_t = f(t) + u_t, \quad (1)$$

где $f(t)$, $t = \overline{1, T}$, и есть тренд, а ее аргумент t интерпретируется как время. Величины $\{u_t\}_{t=1}^T$ в (1) случайны, некоррелированы, с нулевыми математическими ожиданиями и одинаковой ограниченной дисперсией:

$$E\{u_t\} = 0, \quad D\{u_t\} = E\{u_t^2\} = \sigma^2 < +\infty; \quad (2)$$

$$E\{u_t u_l\} = 0, \quad \forall t, l = \overline{1, T}, \quad l \neq t,$$

и имеют смысл ошибок наблюдений.

Заданы $L \geq 2$ различных трендовых моделей. Каждая такая трендовая модель Ω_l определяется своим трендом $f_l(t)$, $t = \overline{1, T}$, $l \in S$, где $S = \{1, \dots, L\}$ – множество номеров этих моделей. Все тренды $\{f_l(\cdot)\}_{l \in S}$ отцентрированы относительно нулевого уровня в том смысле, что

$$\sum_{t=1}^T f_l(t) = 0, \quad l \in S. \quad (3)$$

Модель Ω_l ($l \in S$) содержит разные тренды вида

$$f_l^a(t) = f_l(t) + a, \quad t = \overline{1, T}; \quad (4)$$

где $a \in R$ – параметр сдвига. Таким образом, тренд $f_l(\cdot)$ из (3) является типовым и задает

«тенденцию» [4], определяющую модель Ω_l и не зависящую от абсолютных значений текущего тренда (4), принадлежащего этой модели.

Таким образом, по реализации $X = \{x_t\}_{t=1}^T$, которой соответствует ненаблюдаемый реальный тренд $f(\cdot)$ из (1), необходимо решить, к какой трендовой модели из $\{\Omega_l\}_{l \in S}$ она «ближе». Заранее нужно определить это понятие, а также предложить критерий эффективности принимаемых решений [1].

II. РЕШАЮЩЕЕ ПРАВИЛО И ОПРЕДЕЛЕНИЕ ЕГО РИСКА

Воспользуемся методом наименьших квадратов (МНК) [3] и построим по реализации $X = \{x_t\}_{t=1}^T$ решающее правило (РП) [1, 3] $d = d(X) \in S$, относящее реализацию X к той модели из $\{\Omega_l\}_{l \in S}$, к базовому тренду (3) которой она ближе. РП будет иметь вид:

$$d = d(X) = \arg \min_{l \in S} \min_{a \in R} \sum_{t=1}^T (x_t - f_l^a(t))^2 = \quad (5)$$

$$= \arg \min_{l \in S} \sum_{t=1}^T (x_t - \bar{x}_T - f_l(t))^2,$$

где

$$\bar{x}_T = \frac{1}{T} \sum_{t=1}^T x_t \quad (6)$$

– арифметическое среднее значений отсчетов из реализации $X = \{x_t\}_{t=1}^T$.

Определим множество $D^\circ \subseteq S$ номеров тех моделей из $\{\Omega_l\}_{l \in S}$, к которым реальный тренд $f(\cdot)$ ближе:

$$D^\circ = \left\{ k : \rho_*(f, f_k) = \min_{l \in S} \rho_*(f, f_l) \right\}, \quad (7)$$

где

$$\rho_*^2(f, f_l) = \min_{a \in R} \sum_{t=1}^T (f(t) - f_l^a(t))^2 = \quad (8)$$

$$= \sum_{t=1}^T (f(t) - \bar{f}_T - f_l(t))^2,$$

$$\bar{f}_T = \frac{1}{T} \sum_{t=1}^T f(t), \quad l \in S.$$

В (7), (8) \bar{f}_T имеет смысл значения параметра сдвига, «приводящего» реальный тренд $f(\cdot)$ из (1) к нулевому уровню в смысле (3):

$$\sum_{t=1}^T (f(t) - \bar{f}_T) = 0,$$

а $\rho_*(f, f_l)$ – евклидова расстояния между трендом $f(\cdot)$ реализации X (с учетом сдвига на нулевой уровень) и базовым трендом $f_l(\cdot)$, определяющим модель Ω_l ($l \in S$). В качестве меры эффективности принимаемых решений, будем использовать риск [1]

$$r_T = P\{d(X) \notin D^o\}, \quad (9)$$

имеющий смысл вероятности не отнести при помощи РП $d = d(X) \in S$ из (5), (6) реализацию X к той трендовой модели, к которой она ближе в смысле базовых трендов $\{f_l(\cdot)\}_{l \in S}$ из (3).

Если $D^o = \{d^o\}$ – есть только одна ближайшая к тренду $f(\cdot)$ реализации X модель из $\{\Omega_l\}_{l \in S}$, то

$$r_T = P\{d(X) \neq d^o\}. \quad (10)$$

Чем меньше (ближе к 0) значения риска r_T ($0 \leq r_T \leq 1$) из (9), (10), тем эффективнее принимаемые при помощи РП $d = d(X)$ решения. Отметим также, что РП (5), (6), благодаря сдвигу реализации $X = \{x_t\}_{t=1}^T$, по которой выносится решение, на нулевой уровень:

$$\sum_{t=1}^T (x_t - \bar{x}_T) = 0,$$

позволяет из нескольких ($m \geq 2$) реализаций $X^{(j)} = \{x_t^{(j)}\}_{t=1}^{T_j}$, $j = \overline{1, m}$, отнесенных РП (5), (6) к одной и той же трендовой модели: $d = d(X^{(1)}) = \dots = d(X^{(m)}) \in S$, выбрать наиболее близкую к ней. Ее номер

$$j^* = \arg \min_{1 \leq j \leq m} \sum_{t=1}^{T_j} (x_t^{(j)} - \bar{x}_{T_j}^{(j)} - f_d(t))^2,$$

$$\bar{x}_{T_j}^{(j)} = \frac{1}{T_j} \sum_{t=1}^{T_j} x_t^{(j)}.$$

III. ВЫЧИСЛЕНИЕ РИСКА

Как и в [1], вычислим риск РП $d = d(X)$ в ситуации, когда к тренду $f(\cdot)$ наблюдаемой реализации X наиболее близок лишь один из $L \geq 2$ базовых трендов $\{f_l(\cdot)\}_{l \in S}$, и

$$d^o = \arg \min_{l \in S} \rho_*(f, f_l) \quad (11)$$

– истинный номер ближайшего базового тренда через расстояния (8).

Введем обозначения:

$$\rho(f, f_l) = \sqrt{\sum_{t=1}^T (f(t) - f_l(t))^2}, \quad l \in S;$$

$$\rho(f_l, f_k) = \sqrt{\sum_{t=1}^T (f_l(t) - f_k(t))^2}, \quad l, k \in S,$$

– обычные евклидовы расстояния между соответствующими трендами (без учета сдвига тренда $f(\cdot)$ на нулевой уровень).

Пусть реализация $X = \{x_t\}_{t=1}^T$ и трендовые модели $\{\Omega_l\}_{l \in S}$ определяются соотношениями (1) – (4), а случайные величины $\{u_t\}_{t=1}^T$ в (2), вдобавок, независимы в совокупности и одинаково распределены по нормальному закону $N_1(0, \sigma^2)$ ($0 < \sigma^2 < +\infty$). Если $D^o = \{d^o\}$, где $d^o \in S$ – единственный истинный номер (11) ближайшей к X трендовой модели, то риск r_T из (10) РП (5), (6)

$$r_T = \quad (12)$$

$$= 1 - E \left\{ \prod_{\substack{l \in S \\ l \neq d^o}} U \left(z_l + \frac{\rho^2(f, f_l) - \rho^2(f, f_{d^o})}{2\sigma \rho(f_l, f_{d^o})} \right) \right\}$$

где $U(y) = \{1, y \geq 0; 0, y < 0\}$ – единичная функция Хэвисайда, $\sigma = \sqrt{\sigma^2}$ – среднее квадратическое отклонение, случайные величины z_l , $l \in S$, $l \neq d^o$, имеют стандартное нормальное распределение: $L\{z_l\} = N_1(0, 1)$, а их совместное распределение является многомерным нормальным [3] со следующими ковариациями ($l, k \in S$, $l \neq d^o$, $k \neq d^o$):

$$\text{Cov}\{z_l, z_k\} = \frac{\sum_{t=1}^T (f_l(t) - f_{d^o}(t))(f_k(t) - f_{d^o}(t))}{\rho(f_l, f_{d^o})\rho(f_k, f_{d^o})}.$$

Полученное соотношение (12) позволяет вычислить риск r_T РП $d = d(X)$ и оценить теоретически его эффективность. Однако, как и в [1], простой вид риск имеет лишь при $L = 2$:

$$r_T = \Phi \left(- \frac{|\rho^2(f, f_1) - \rho^2(f, f_2)|}{2\sigma \rho(f_1, f_2)} \right),$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{w^2}{2}\right) dw$, $z \in R$ – функция распределения стандартного нормального закона.

1. Жук, Е. Е. Статистическое отнесение реализаций нестационарных временных рядов к заданным трендовым моделям / Е. Е. Жук // Вес. Нац. акад. наук Беларусі, Сер. фіз – мат. навук, – 2017. – № 2. – С. 5259.
2. Андерсон, Т. Статистический анализ временных рядов: пер. с англ. / Т. Андерсон. – М.: Мир, 1976. – 759 с.
3. Харин, Ю. С. Математическая и прикладная статистика / Ю. С. Харин, Е. Е. Жук. – Минск: БГУ, 2005. – 276 с.
4. Элдер, А. Как играть и выигрывать на бирже / А. Элдер. – М.: Альпина Паблишер, 2017. – 472 с.

Содержание

Пленарные доклады	7
Таранчук В. Б. «Новые компьютерные технологии, анализ и интерпретация геоданных»	8
Голенков В. В., Гулякина Н. А., Шункевич Д. В. «Текущее состояние и направления развития технологий искусственного интеллекта»	11
Пархименко В. А. «О некоторых вопросах алгоритмического маркетинга»	17
Секция «Системы управления»	21
Акинфина М. А., Бондаренко С. П. «Система анализа и оценки рисков страхования»	22
Акиншева И. В., Пантюхов В. А. «Управление запросами к базе данных пользователей социальной сети через паттерн MVC»	24
Бранцевич П. Ю., Деменковец Д. В. «Система отображения информации на базе контроллера STM и платформы Android»	26
Черемисинов Д. И. «Метод логико-временного анализа цифровых систем с ограничением на время отклика»	28
Черемисинова Л. Д. «К анализу параллельного управления на основе расширенного графа достижимости»	30
Гинько Д. В., Цингель А. А., Курулев А. П. «Система автоматического управления в автомобилях»	32
Дорошко А. С. «Исследование эффективности внедрения АСУ ТП на химических предприятиях в Республике Беларусь»	34
Elzein I., Petrenko Y. N. «Improving the P&O and Incremental Conductance algorithms operations in Photovoltaic station through a Model Predictive Control»	36
Гринюк Д. А. и др. «Эффективность линейного и нелинейного 2DoF регуляторов»	38
Илюшин И. Э., Кожевников М. М. «Автоматическое управление роботами-манипуляторами для дуговой сварки с учетом технологических ограничений»	40
Карпович С. Е., Зубов Г. А., Войтов А. Ю. «Системы перемещений в прецизионном оборудовании микроэлектроники»	42
Корчажкина О. М. «Модель онтологии понятийных категорий планиметрии»	44
Кулаков Г. Т., Кухоренко А. Н. «Инвариантная система автоматического управления уровнем воды в барабане парового котла без измерения внешних возмущений»	46
Давыдов М. В. и др. «Методика оценки и программной коррекции неравномерности АЧХ ВЧ-тракта генератора для широкополосного перестраиваемого акустооптического фильтра»	48
Новицкий И. О., Шилин Л. Ю. «Метод биоэлектрографии на современном этапе»	50
Олиферович Н. М. и др. «Апериодический цифровой регулятор с прогнозируемым уровнем сигнала управления»	52
Полудворянин С. М., Нестеренков С. Н. «Проблемы безопасности sip-телефонии»	54
Пригара В. Н. «Процессы истечения и сводообразования в бункерах»	56
Пригара В. Н. «Обзор и сравнение существующих методов очистки поверхностей»	58
Симаньков В. И. «Устранение конфликтов маршрутов при облете целей группой БПЛА»	60

Ситкевич Т. А. «Разработка устройства для автоматической намотки трансформатора»	62
Соколов В. И., Хаджинов М. К. «Вопросы быстродействия в системах управления электроприводом»	64
Сорокин А. А. «Система сбора информации о состоянии зоны покрытия сети для управления развитием инфраструктуры операторов связи»	66
Струц Д. А. «Оптимизация времени ожидания при диспетчеризации группы лифтов на основе генетического алгоритма»	68
Свинарский М. В. и др. «Синтез устройства совместного измерения-распознавания с адаптацией к углам пространственной ориентации радиолокационных объектов»	70
Тимахович А. Е., Хаджинов М. К. «Модернизация программного комплекса промышленного сервопривода»	72
Титко Д. С., Салманзадех Г. Й., Марко А. Ф. «Система перемещений установки автоматического контроля оригиналов топологии»	74
Волков А. В. и др. «Анализ и разработка модулей для групповой робототехники»	76
Янцевич М. А., Дубовик И. А. «Преимущества и особенности использования неминимально-фазового коэффициента отражения в задачах широкополосного согласования сопротивлений»	78
Ярохович А. А., Хаджинов М. К. «Проблемы систем управления турбогенераторной установкой»	80
Заблоцкий В. В., Рудикова Л. В. «О разработке системы рекомендаций и прокладки маршрутов к туристическим объектам»	82
Жилинская Е.Р., Кардаш И.П., Захарьев В.А. «Применение технологии реестра блоков транзакций для построения распределенных систем»	84
Секция «Интеллектуальные информационные технологии» ...	87
Аксёничков А. А. «Интеллектуальная программа имитационного моделирования эксплуатационной работы на пассажирской железнодорожной станции»	88
Бабеня В. И. «Определение злокачественных опухолей грудной железы на гистологических снимках»	90
Бобков А. В. «Компонент визуализации графоподобных диаграмм на основе их формального описания в базе знаний»	92
Доронин А. К., Липницкий В. А. «Предсказательная модель машинного обучения для решения задачи классификации уязвимостей компьютерных систем»	94
Дюбина Е. А., Давыденко И. Т. «Средства верификации машин обработки знаний интеллектуальных систем»	96
Астахова Н. М., Федорович А. Г. «Онтология предметной области поддержки, подготовки и проведения конференции»	98
Гончаров И. В., Потешкин А. С., Курулев А. П. «Архитектура современной нейронной сети для сегментации объектов на изображениях»	100
Гулякина Н.А., Давыденко И.Т., Зверуго А.В. «Актуальные проблемы в области разработки гибридных баз знаний»	102
Губарева О. И. «Автоматизированная система поиска элементов позвонка на КТ-изображениях с использованием анализа контуров»	104
Ивашенко В. П. «Модель обработки строк и списков данных для систем, управляемых знаниями»	106

Ивашенко В. П. «Идентификация бинарно порождаемых событий процессов обработки знаний для семантического протоколирования»	108
Ивашенко В. П. «Онтологическое моделирование пространственно-временных отношений» ..	110
Карпук А. А. «Новые задачи управления использованием радиочастотного спектра и пути их решения»	112
Кочурко А. А., Абламейко С. В. «Выделение основной дорожной сети на изображениях земной поверхности с помощью нейронной сети U-Net»	114
Куручка К. С., Оныськив В. В. «Программный комплекс для управления роботом из виртуальной реальности»	116
Лапицкая Н. В. и др. «Смартфон как инструмент предоставления персональных рекомендаций стратегии сбалансированного питания»	118
Левко С. В., Шадраков Р. Н. «Байесовские нейронные сети»	120
Марко А. Ф., Чеушев К. В., Лобашинский М. В. «Программное средство для обеспечения целостности при разработке и эксплуатации системы автоматизированного управления транспортным оборудованием»	122
Нестеренков С. Н., Белов К. П. «Использование генетического алгоритма для нахождения коэффициентов и структуры нейронной сети»	124
Нестеренков С. Н., Видничук В. Н., Шинкевич Н. Н. «Система определения вероятности успешного прохождения собеседования соискателем на основе экспертных оценок»	126
Пилипчук Л. А., Лагуто А. А., Пилипчук А. С. «Об аналитических и численных решениях двух типов разреженных систем линейных алгебраических уравнений»	128
Пилипчук А. С. и др. «О методах и технологиях моделирования процессов экологии и визуализации пространственных данных»	130
Родионова О. С. «Средства информационной поддержки разработчиков интеллектуальных систем на основе открытой семантической технологии»	132
Шадраков Р. Н., Левко С. В. «Использование нейронных сетей для поиска дорожных знаков»	134
Шараев Е. В. «Оптимизация запросов к базе данных с помощью алгоритмов искусственного интеллекта»	136
Шинкевич Н. Н. «Оценка релевантности контента в режиме реального времени»	138
Таранчук В. Б. «Средства и примеры интеллектуальной обработки данных геологических моделей»	140
Вешторт А. В. «Автоматизированный анализ вредоносного программного обеспечения с применением рекуррентных нейронных сетей»	142
Жиленкова В. В. «Создание многовариантных тестовых заданий средствами Wolfram Mathematica»	144
Секция «Проектирование встраиваемых систем»	147
Бибило П. Н., Кириенко Н. А., Ланкевич Ю. Ю. «Исследование эффективности оптимизации представлений систем булевых функций на основе блочного разбиения и разложения Шеннона»	148
Бибило П. Н., Романов В. И. «Использование скриптов для логической оптимизации и синтеза логических схем»	150
Дубовик И. А. «Решение задачи синтеза частотно-избирательной цепи с оптимальными характеристиками с помощью алгоритма оптимального синтеза по заданному критерию»	152

Емельянов Д. В. «Система сбора и анализа данных для приложений реального времени» ...	154
Иванюк А. А. «Особенности реализации симметричных путей ФНФ типа арбитр на ПЛИС»	156
Качинский М. В., Станкевич А. В. «Конвейерный процессор хэш-функции SHA-256»	158
Кардаш С. Н. «Экспериментальное исследование алгоритмов дизъюнктивно-конъюнктивного разложения системы полностью определенных булевых функций»	160
Клыбик В. П., Иванюк А. А. «Перспективные возможности обеспечения безопасности инфраструктуры IoT»	162
Королёв Н. А. «Использование библиотек LIBERTY при построении математических моделей для оценки энергопотребления КМОП-микросхем»	164
Костюк Д. А., Латий О. О., Маркина А. А. «Биометрическая измерительная система для оценки состояния пользователя ПК»	166
Латий О. О. «Программное обеспечение для визуализации потока данных, поступающих в ПЭВМ от микроконтроллерной системы»	168
Секция «Математическое моделирование и компьютерная графика»	171
Аникин В. И., Козырин Д. И. «Сравнительное моделирование марковских цепей в среде программирования R и Microsoft Excel»	172
Ключникова К. С., Бычков П. А. «Образовательная игра «Котёнок Звуковкин»»	174
Исаев В. О., Свириденко А. А., Шанчук Р. А. «Алгоритм аппроксимации частотных характеристик СВЧ транзисторов, представленных в численном виде, на дискретном ряде частот»	176
Курочка К. С., Комракова Е. В. «Расчет термоупругих деформаций в сэндвич-панели» ...	178
Нестеренков С. Н. и др. «Эргономичный пользовательский интерфейс как средство повышения эффективности работы автоматизированной системы управления в учреждении высшего образования»	180
Шамына А. Ю., Ардяко А. Д. «Программное средство для моделирования распространения радионуклидов в окружающей среде на основе гауссовой модели»	182
Секция «Системы обработки информации»	185
Assanovich V. A., Veretilo Yu. N. «Fuzzy Secure Sketch Biometric Scheme Based on Non-Binary Turbo Codes»	186
Рудикова Л. В., Бандысик С. Ю. «Использование технологии складирования данных для построения архитектуры системы сбора и анализа данных произведений исторической ценности»	188
Боброва Т.С., Кузнецова О.В. «Применение облачных технологий при разработке медицинских диагностических аппаратно-программных комплексов»	190
Буй П. М., Кульгавик С. Г. «Методика перекрестной оценки уязвимостей и угроз безопасности информационных систем железнодорожного транспорта»	192
Бухаров Д. С. «К вопросу автоматизации настройки АЛАР КПАМ: фиксация изменения сопротивления»	194
Бухаров Д. С. «К вопросу автоматизации настройки АЛАР КПАМ: фиксация колебания тока»	196
Быстрова М. В., Козадаев И. А. «Классификация вакансий с целью последующей оптимизации публикации объявлений»	198

Чепелева М. К., Яцков Н. Н., Назаров П. В. «Исследование статистической устойчивости дифференциально выраженных генов и обогащенных биологических функций раковых клеток»	200
Досин А. Н., Недзьведь А. М., Белоцерковский А. М. «Построение поверхности изображения по произвольным линиям»	202
Дрозд П. С., Козадаев К. В. «Методы обработки пользовательских данных для прогнозирования бизнес-процессов»	204
Канунников И. С., Парамонов А. И. «Влияние показателей игры теннисиста на исход матча»	206
Климук И.В., Свидрицкий А.С., Яцков Н.Н. «Разработка алгоритмов анализа биочипов ДНК с учетом весовых факторов качества спотов»	208
Козарь Р. В. и др. «Разработка эвристического алгоритма поиска оптимального маршрута»	210
Кузнецов А. В., Рудикова Л. В., Жавнерко Е. В. «Об основных аспектах разработки интернет-приложения «Виртуальный помощник туриста»»	212
Кузнецова О. В., Протченко Е. В., Боброва Т. С. «Применение нейронных сетей для решения задач медицинской диагностики»	214
Липницкий В. А., Сергей А. И. «Оптимизация вычисления двумерных рекуррентных сверток при помощи быстрого преобразования Фурье»	216
Лобач С. В. «Тестирование скачков изменения характеристик сигналов с помощью вейвлетов»	218
Лобач В. И. «Вейвлет-анализ в исследовании неоднородностей случайных последовательностей»	220
Лутковский В. М., Волков А. В. «Модели адаптивного обучения человека и машины»	222
Меркулов Р. И., Лобач В. И. «Статистический анализ панельных данных на основе моделей в пространстве состояний»	224
Mikhailovskaya L. V., Valakhanovich E. V. «On the classification of information security threats»	226
Моженкова Е. В., Парамонов А. И. «Модель представления информации в корпоративных информационных системах»	228
Мысливец О. Р., Рудикова Л. В. «Построение хранилища данных для задач обработки данных пользователей социальных сетей»	230
Нефёдов Д. С., Курилович В. В., Дубков А. В. «Математическая модель возмущений напряженности электрического поля, создаваемого вертолетом»	232
Неживинская А. Ю., Рудикова Л. В. «О разработке системы обработки данных о парфюмерных брендах»	234
Поздняков С. В., Давыдов М. В. «Информационная система моделирования техногенных катастроф»	236
Потараев В. В. «Применение нейронной сети для обработки текстовой информации»	238
Ревотюк М. П., Гибулина Е. М., Бруй Н. М. «Распознавание проблемных ситуаций в системах координации взаимодействующих агентов»	240
Ревотюк М. П., Грабовский Д. В. «Реоптимизация планов управления ресурсами в потоках работ»	242
Рогачёв Н. Е. «Использование нейронных сетей-классификаторов для обнаружения изменений на спутниковых снимках»	244

Рудикова Л. В., Жавнерко Е. В. «О разработке системы хранения и обработки данных произведений художественной ценности с использованием технологии складирования данных»	246
Садов С. В. «Алгоритм формирования метрики лица в задаче определения эмоционального состояния человека по видеопоследовательности лица человека»	248
Сасин Е. А., Сидорович А. С. «Оценка вычислительной сложности обобщенного алгоритма кластеризации»	250
Сасин Е. А., Сидорович А. С. «Анализ баз изображений сетчатки глаза человека»	252
Татур М. М., Сасин Е. А., Сидорович А. С. «Моделирование системы управления торможением мобильной роботизированной платформы»	254
Савчук А. А., Нестеренков С. Н. «Автоматизация контроля знаний как метод оптимизации процесса обучения»	256
Захарченко К. В., Шилин Л. Ю. «Схема офлайн-генерации эфемерных ключей шифрования в одноранговой сети»	258
Шлеменков А. А., Гусак Я. О. «Использование многоуровневой модели для эффективного управления дамбой и предсказания наводнений»	260
Шульга Е. С., Филипук Д. С., Сурков К. А. «Методика создания возобновляемых рабочих процессов в автоматизированных системах»	262
Шуляк Е. В., Гуринович А. Б., Смирнов В. Л. «Прогнозные задачи в CRM-системах»	264
Сивко Б. В. «Автоматизация верификации программного обеспечения связанных с безопасностью систем»	266
Стригалёв Л. С., Тиханович Т. В., Кузнецова О. В. «Технологии искусственного интеллекта в образовании»	268
Таранчук В. Б., Таранчук В. В. «Методы и инструменты оценки точности компьютерных геологических моделей»	270
Ткачев И. Е., Рудикова Л. В. «О разработке приложения для оптимизации работы логистических компаний на базе мобильной платформы iOS»	272
Тюшкевич М. В. «Использование рекуррентных нейронных сетей при распознавании и анализе состояний динамических объектов»	274
Волков А. В., Яцков Н. Н., Гринев В. В. «Анализ алгоритмов выбора признаков экзонов»	276
Волков А. В., Яцков Н. Н., Гринев В. В. «Имитационная модель для тестирования алгоритмов отбора признаков»	278
Якубович А.В., Лежнёв А.В. «Идентификация участников общественных мероприятий на основе данных социальных сетей»	280
Яцков Н. Н., Скакун В. В., Гринев В. В. «Программный пакет RNAexploreR для предсказания вариантов альтернативного сплайсинга в первичных мРНК химерного онкогена RUNX1/RUNX1T1 человека»	282
Яцков Н. Н., Скакун В. В., Гринев В. В. «Статистический подход к предсказанию событий альтернативного сплайсинга в первичных мРНК генов человека»	284
Заерко Д. В., Липницкий В. А. «Применение модифицированных алгоритмов Java ANPR для автоматического распознавания номеров автомобилей»	286
Заграй В. Ю., Гуринович А. Б. «Фильтрация спам-писем с помощью алгоритмов на основе нейронных сетей»	288
Зайко Е. В. и др. «Использование яркостных признаков изображений в интересах распознавания воздушных объектов в оптико-локационных системах»	290

Жук Е. Е., Чернявский Д. В. «Определение трендовой модели типа сдвига для нестационарного временного ряда»292

Научное издание

**Информационные технологии и системы 2018
(ИТС 2018)**

Материалы международной научной конференции
(БГУИР, Минск, Беларусь, 25 октября 2018)

**Information Technologies and Systems 2018
(ITS 2018)**

Proceeding of The International Conference
(BSUIR, Minsk, Belarus, 25th October 2018)

В авторской редакции

Ответственный за выпуск *Л.Ю. Шилин, А.Б. Гуринович*

Компьютерная верстка *Н.С. Жилач, А.Ф. Трофимович*

Дизайн обложки *А.А. Навроцкий*

Подписано в печать 18.10.2017. Формат 60x84 1/8. Бумага офсетная. Гарнитура «BookAntiqua».
Отпечатано на ризографе. Усл. печ. л. 41,15. Уч.-изд. л. 45,3. Тираж 100 экз. Заказ 401.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
ЛИ №02330/0494371 от 16.03.2009. ЛП №02330/0494175 от 03.04.2009.
220013, Минск, П. Бровки, 6.