

Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»



**Информационные технологии и системы 2023
(ИТС 2023)**

Материалы международной научной конференции
(БГУИР, Минск, Беларусь, 22 ноября 2023)

Information Technologies and Systems 2023
(ITS 2023)

Proceeding of the International Conference
(BSUIR, Minsk, Belarus, 22nd November 2023)

Минск БГУИР 2023

УДК 004
ББК 32.81
И74

Редакционная коллегия:

Л. Ю. Шилин (главный редактор), А. А. Иванюк, С. В. Колосов, В. С. Муха,
В. В. Голенков, А. В. Марков, Д. П. Кукин, А. А. Навроцкий,
В. И. Журавлев, А. Ф. Трофимович, Н. В. Гракова, Д. В. Шункевич,
М. В. Ковалёв (ответственный секретарь)

Информационные технологии и системы 2023 (ИТС 2023) =
И74 Information Tehnologies and Systems 2023 (ITS 2023) : материалы
международной научной конференции, Минск, Беларусь, 22 ноября / Л. Ю.
Шилин [и др.]. – Минск : БГУИР, 2023.
ISBN

Сборник включает прошедшие рецензирование доклады международной научной конференции «Информационные технологии и системы 2023» (ИТС 2023).

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 32.973.202

ISBN

«Белорусский государственный
университет
информатики и радиоэлектроники», 2023

ПРЕДИСЛОВИЕ

22 ноября 2023 года в городе Минске прошла международная научная конференция «Информационные технологии и системы» (ИТС 2023, 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 98 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 «Автоматизированные 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 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, Республики Таджикистан.

Докладчики представляли такие высшие учебные заведения как: Белорусский государственный университет информатики и радиоэлектроники, Институт информационных технологий, Национальный детский технопарк, Белорусский государственный университет, Академия управления при Президенте Республики Беларусь, Минский государственный лингвистический университет, Военная академия Республики Беларусь, Гомельский государственный технический университет имени П. О. Сухого, Полоцкий государственный университет имени Ефросинии Полоцкой, МИРЭА – Российский технологический университет, Институт МПСУ, Национальный исследовательский университет, «Московский институт электронной техники», Санкт-Петербургский государственный университет, Московский институт электронной техники, Таджикский технический университет имени М. Осими.

Кроме того, в работе конференции приняли участие представители таких организаций как ЗАО «Инженерный центр Ядро» (Минск, РБ), ООО «Ядро микропроцессоры» (Москва, Россия).

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
PhD	Maksim Davydov	First Vice-Rector of BSUIR
PhD	Victor Rybak	Vice-Rector for Education of BSUIR
PhD	Evgeni Shneiderov	Vice-Rector for Education of BSUIR
PhD	Viktor Stempitsky	Vice-Rector for Research of BSUIR
	Dmitry Kuznetsov	Vice-rector for Personal Development of BSUIR
Prof.Dr.	Leonid Shilin	Dean of The Faculty of Information Technologies and Control of BSUIR

PROGRAMME COMMITTEE

Prof.Dr.	Vladimir Mukha	Professor of Information Technologies in Automated Systems Department of BSUIR
Prof.Dr.	Vladimir Golenkov	Professor of Intelligent Information Technologies Department of BSUIR
Prof.Dr.	Vladimir Golovko	Head of Intelligent Information Technologies Department of BrSTU
Prof.Dr.	Alexander Ivaniuk	Professor of Computer Science Department of BSUIR
Prof.Dr.	Stanislav Kolosov	Professor of Computational Methods and Programming Department of BSUIR
PhD	Dmitry Kukin	Head of Computational Methods and Programming Department of BSUIR
PhD	Alexander Markov	Head of Control Systems Department of BSUIR
PhD	Anatoly Navrotsky	Head of Information Technologies in Automated Systems Department of BSUIR
PhD	Daniil Shunkevich	Head of Intelligent Information Technologies Department of BSUIR
PhD	Vadim Zhuravliov	Head of Fundamental Electrical Engineering Department of BSUIR

Пленарные доклады

Plenary session

КЛЮЧЕВЫЕ ПРОБЛЕМЫ И СТРАТЕГИЧЕСКИЕ ЦЕЛИ РАБОТ В ОБЛАСТИ 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, guliakina, shunkevich}@bsuir.by

В работе рассмотрены основные методологические проблемы текущего состояния работ в области искусственного интеллекта, а также ключевые задачи и стратегические цели развития указанных работ. Обоснована необходимость перехода к согласованной комплексной автоматизации человеческой деятельности.

I. ПРОБЛЕМЫ ТЕКУЩЕГО СОСТОЯНИЯ РАБОТ В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

К методологическим проблемам текущего состояния *Деятельности в области искусственного интеллекта* относятся следующие:

- Отсутствует общая формальная теория интеллектуальных компьютерных систем
:= [Отсутствует конвергенция (сближение) и интеграция различных направлений искусственного интеллекта]

⇒ *подпроблема**:

[Отсутствует общая формальная теория интероперабельных интеллектуальных компьютерных систем]

⇒ *подпроблема**:

[Отсутствует унификация и стандартизация интероперабельных интеллектуальных компьютерных систем]

- Отсутствует комплексная технология проектирования, производства и модернизации интеллектуальных компьютерных систем
- Отсутствует кадровый потенциал, необходимый для создания общей теории интеллектуальных компьютерных систем и комплексной технологии их проектирования, производства и модернизации
- У современных специалистов в области искусственного интеллекта отсутствует необходимый уровень интероперабельности, мотивации к конвергенции и интеграции различных направлений искусственного интеллекта

II. ГЛАВНЫЙ ПРОДУКТ ДЕЯТЕЛЬНОСТИ В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Деятельность в области искусственного интеллекта

⇒ *главная стратегическая цель (главный продукт)**:

Комплексная автоматизация всех видов и областей человеческой деятельности на основе массового использования интеллектуальных компьютерных систем
:=

[Глобальная сеть (экосистема), состоящая из семантически совместимых интероперабельных интеллектуальных компьютерных систем, самостоятельно и осмысленно взаимодействующих как между собой, так и с людьми]

⇒ *подцель**:

- Стратегическая научно-техническая программа комплексной автоматизации всех видов и областей человеческой деятельности на основе массового использования интероперабельных интеллектуальных компьютерных систем
- Формализация и конвергенция различных видов и областей человеческой деятельности

Указанный подход к комплексной автоматизации человеческой деятельности означает переход на принципиально новый технологический уклад и уровень развития используемых человеком технологий.

III. АНАЛИЗ ТЕКУЩЕГО СОСТОЯНИЯ АВТОМАТИЗАЦИИ ЧЕЛОВЕЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Основная проблема текущего состояния *Комплекса средств автоматизации человеческой деятельности* заключается в том, что доминирует независимая локальная автоматизация различных видов и областей человеческой деятельности, т. е. отсутствует комплексный (интегральный) подход к автоматизации человеческой деятельности в целом.

Подпроблемами указанной проблемы являются следующие:

- Каждый человек или коллектив должен вручную декомпозировать решаемые им комплексные задачи, требующие использования различных локальных средств автоматизации, а также должен вручную осуществлять взаимодействие этих средств при решении указанных комплексных задач.
- Каждый человек или коллектив должен знать, какие потенциально полезные ему

средства автоматизации существуют в текущий момент и как с этими средствами вручную взаимодействовать с помощью соответствующих пользовательских интерфейсов (т. е. знать язык общения с этими средствами и возможности этих средств).

- Отсутствует унификация пользовательских интерфейсов для различных локальных средств автоматизации (даже для функционально эквивалентных средств), т. е. с разными даже эквивалентными средствами автоматизации необходимо общаться на разных «языках».

Следствием того, что организация управления различными локальными средствами автоматизации носит эклектичный характер (разные средства управляются по-разному), являются большие затраты на формирование у пользователей навыков управления различными средствами автоматизации и недостаточно полное использование всех возможностей применяемых средств.

Современный этап информатизации и цифровой трансформации человеческой деятельности и соответствующие требования к компьютерной грамотности пользователей носят фактически издевательский характер, особенно для неподготовленных пользователей. Пользователя заставляют знать не только смысл своей задачи, но и то, как эта задача решается с помощью компьютерных систем. Причем в разных аналогичных компьютерных системах это делается по-разному.

- С расширением многообразия различных локальных средств автоматизации человеческой деятельности существенно возрастают накладные расходы каждого человека, а также каждого коллектива на организацию взаимодействия с этими средствами автоматизации и, соответственно, существенно сокращается доля времени, которую каждый человек или коллектив может потратить на принципиально неавтоматизируемую (в первую очередь творческую, созидательную) деятельность.

В настоящее время при расширении многообразия средств автоматизации локальных видов и областей человеческой деятельности человек незаметно для себя становится придатком этих средств, уровень квалификации которого определяется не уровнем понимания смысла выполняемой деятельности, а умением управлять соответствующими средствами автоматизации, т. е. умением грамотно нажимать на соответствующие клавиши, кнопки, рычаги.

Если провести аналогию современного характера использования средств автоматизации

человеческой деятельности с деятельностью в области музыкального творчества, то можно сказать, что на современном уровне использования средств автоматизации доминируют не творцы, не композиторы, а исполнители, которые неспособны внести собственный вклад в исполняемое произведение, вклад, раскрывающий замысел композитора, и которые тем более неспособны на уместную и гармоничную импровизацию. Очевидно, что такое положение дел существенно снижает темпы технологического прогресса, а также темпы прогресса во всех автоматизируемых областях человеческой деятельности.

Поскольку на современном этапе развития средств автоматизации человеческой деятельности доминирующий характер имеет информационная (умственная) деятельность (в том числе и деятельность по управлению различными средствами автоматизации физической деятельности), создаются хорошие предпосылки для унификации принципов организации управления этими средствами.

IV. АКТУАЛЬНОСТЬ ПЕРЕХОДА ОТ ЛОКАЛЬНОЙ К КОМПЛЕКСНОЙ АВТОМАТИЗАЦИИ ЧЕЛОВЕЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Прежде чем приступить к рассмотрению актуальности перехода от локальной к комплексной автоматизации человеческой деятельности, подчеркнем, что принципиальная возможность реализации такого перехода возникла сравнительно недавно. Это связано с тем, что указанный переход на новый уровень автоматизации человеческой деятельности предполагает автоматизацию не всегда априори предсказуемого и предусматриваемого взаимодействия между различными локальными средствами автоматизации и поэтому требует использования не просто интеллектуальных компьютерных систем, а интероперабельных интеллектуальных компьютерных систем, способных к самостоятельному эффективному взаимодействию в процессе коллективного решения комплексных задач, теория и технология разработки которых только недавно получила свое развитие.

Человечество достигло больших успехов в автоматизации локальных видов и областей человеческой деятельности. При этом человек превратился в субъекта, не только управляющего данными средствами, но и в субъекта, связующего их между собой. Для того чтобы в этом убедиться, достаточно проанализировать, на что тратится время каждого из нас и можно ли это автоматизировать. А время – наш главный и при этом невозобновляемый ресурс. Необходимость именно комплексного подхода к автоматизации человеческой деятельности обусловлена следующим:

- Все виды и области человеческой деятельности взаимосвязаны.
- Автоматизация только локализованных видов и областей человеческой деятельности без автоматизации взаимодействия (связей) вынуждает людей «вручную» выполнять роль связующих звеньев между различными средствами автоматизации.

Массовое внедрение интеллектуальных компьютерных систем во все сферы человеческой деятельности и обеспечение их эффективного взаимодействия для коллективного решения сложных задач «на стыках» различных видов и областей человеческой деятельности переводит автоматизацию этой деятельности на качественно новый уровень.

Локальное использование интеллектуальных компьютерных систем принципиально не может существенно повысить общий уровень автоматизации. Снижение такого уровня происходит на «стыках». Это аналогично тому, что ремонт любой дороги на нескольких локальных участках при сохранении плохого качества дороги на промежутках между этими участками не может существенно повысить пропускную способность ремонтируемой дороги.

Для решения проблемы осознанной (осмысленной) автоматической стыковки средств автоматизации различных локальных видов и областей автоматизируемой человеческой деятельности необходим высокий уровень *интероперабельности* указанных взаимодействующих средств автоматизации (т. е. высокий уровень *семантической совместимости, договороспособности и способности к координации действий*).

V. КОНВЕРГЕНЦИЯ КАК МЕТОДОЛОГИЧЕСКАЯ ОСНОВА КОМПЛЕКСНОЙ АВТОМАТИЗАЦИИ ЧЕЛОВЕЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Зачем нам из-за методологической и системной безграмотности свои-ми руками усложнять создаваемую нами искусственную (рукотворную) часть окружающей среды? Мир, в котором мы живем, и так сложен. Необходимо находить как можно больше аналогий в том, что и как мы делаем, и реализовывать аналогичные компоненты нашей деятельности одинаковым образом. В этом состоит суть конвергенции. Все многообразие видов и областей человеческой деятельности должно быть переосмыслено на предмет их максимально возможной конвергенции для минимизации многообразия (дублирования) технических решений при разработке различных интеллектуальных компьютерных систем автоматизации человеческой деятельности.

Для решения проблемы комплексной автоматизации человеческой деятельности необходима конвергенция (обеспечение совместимости) и глубокая интеграция моделей, методов и средств автоматизации различных видов и областей че-

ловеческой деятельности. В этом как раз и заключается суть интероперабельных интеллектуальных компьютерных систем. Однако указанная конвергенция моделей, методов и средств автоматизации требует также конвергенции самих видов и областей человеческой деятельности (это также существенно сократит многообразие средств автоматизации).

Конвергенция различных видов и областей человеческой деятельности предполагает построение общей формальной теории человеческой деятельности, являющейся основой для эффективной комплексной автоматизации всевозможных видов и областей человеческой деятельности, в основе которой лежит выявление аналогий и минимизация дублирования (многообразия), унификация и стандартизация технических решений. В результате этого многообразие частных (локальных) технических решений должно превратиться в систему семантически совместимых и эффективно взаимодействующих технических систем.

Для того чтобы строго говорить о конвергенции различных видов и областей человеческой деятельности, необходимо построить их четкую иерархическую систему. В качестве примера фрагмента такой системы рассмотрим структуру деятельности в области искусственного интеллекта.

Деятельность в области искусственного интеллекта

:= [Деятельность, направленная на создание интеллектуальных компьютерных систем (искусственных кибернетических систем)]

⇒ *декомпозиция**:

- Научно-исследовательская деятельность в области искусственного интеллекта, направленная на построение теории интеллектуальных компьютерных систем
- Разработка комплекса технологий искусственного интеллекта

:= [Разработка комплекса технологий поддержки жизненного цикла интеллектуальных компьютерных систем]

⇒ *декомпозиция**:

- Разработка и модернизация технологии проектирования интеллектуальных компьютерных систем
- Разработка и модернизация технологии производства интеллектуальных компьютерных систем
- Разработка и модернизация технологии эксплуатации интеллектуальных компьютерных систем
- Разработка и модернизация технологии модернизации

интеллектуальных компьютерных систем

- Инженерная деятельность в области Искусственного интеллекта
:= [Эксплуатация комплекса технологий искусственного интеллекта]
⇒ *декомпозиция**:
 - Проектирование интеллектуальных компьютерных систем (в том числе и глобальной экосистемы интеллектуальных компьютерных систем)
 - Производство интеллектуальных компьютерных систем
 - Эксплуатация интеллектуальных компьютерных систем
 - Модернизация интеллектуальных компьютерных систем
- Подготовка кадров в области искусственного интеллекта
- Общая организация деятельности в области искусственного интеллекта
:= [Разработка стратегии и тактики развития деятельности в области искусственного интеллекта и ее материально-техническое и финансовое обеспечение]

Очевидно, что подобную структуру имеют все области человеческой деятельности, направленные на создание различных сложных искусственных (в том числе и социотехнических) систем. При этом можно говорить об общих принципах организации и поддержки (автоматизации) с помощью интеллектуальных компьютерных систем:

- научно-исследовательской деятельности в самых разных областях;
- проектной деятельности в различных областях;
- производственной деятельности в различных областях;
- подготовки кадров в различных областях.

В основе указанной конвергенции различных видов и областей человеческой деятельности лежит борьба с информационным кризисом (информационным «потопом»), с дублированием накапливаемых человечеством знаний и навыков. Для этого необходим переход на смысловое представление этих знаний, удобное как для людей, так и для интеллектуальных компьютерных систем, а также конвергенция онтологий различных предметных областей (конвергенция соответствующих систем понятий).

Особо следует подчеркнуть, что непосредственная текущая деятельность администраторов (менеджеров) различного уровня по контролю и управлению бизнес-процессами соответствующих им организаций должна быть полностью формализована, автоматизирована и заменена на деятельность по эволюции (модернизации, раз-

витию, совершенствованию) этих организаций. Текучка не должна заедать – она должна просто исчезнуть!

VI. АРХИТЕКТУРА ГЛОБАЛЬНОЙ ЭКОСИСТЕМЫ ИНТЕРОПЕРАБЕЛЬНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ КОМПЬЮТЕРНЫХ СИСТЕМ

Глобальная экосистема интероперабельных интеллектуальных компьютерных систем представляет собой сеть взаимодействующих индивидуальных интеллектуальных компьютерных систем.

Детализация архитектуры экосистемы прежде всего предполагает ролевую типологию (специализацию) индивидуальных интеллектуальных компьютерных систем, входящих в экосистему:

- персональные ассистенты;
- порталы научно-технических знаний;
- интеллектуальные автоматизированные системы управления сложными объектами различного вида (производственными предприятиями, городами, организациями, ...);
- интеллектуальные обучающие системы;
- интеллектуальные системы автоматизированного проектирования (в том числе коллективного) сложных систем различного вида.

Кроме того, можно говорить о типологии индивидуальных интеллектуальных компьютерных систем по характеру их внешней среды:

- индивидуальная интеллектуальная компьютерная система, взаимодействующая только с другими индивидуальными интеллектуальными компьютерными системами;
- корпоративная интеллектуальная компьютерная система, обеспечивающая координацию деятельности соответствующего коллектива интеллектуальных компьютерных систем;
- индивидуальная интеллектуальная компьютерная система, взаимодействующая с другими индивидуальными интеллектуальными компьютерными системами и с физической окружающей средой;
- индивидуальная интеллектуальная компьютерная система, взаимодействующая с другими индивидуальными интеллектуальными компьютерными системами и с конкретным обслуживаемым пользователем.
:= [персональный ассистент пользователя]

VII. НАПРАВЛЕНИЯ САМОСТОЯТЕЛЬНОЙ И ИСКУССТВЕННО ПОДДЕРЖИВАЕМОЙ ЭВОЛЮЦИИ ГЛОБАЛЬНОЙ ЭКОСИСТЕМЫ ИНТЕРОПЕРАБЕЛЬНЫХ ИНТЕЛЛЕКТУАЛЬНЫХ КОМПЬЮТЕРНЫХ СИСТЕМ

Важнейшим фактором, определяющим качество любой кибернетической системы (в том числе и глобальной экосистемы интероперабельных интеллектуальных компьютерных систем),

является не только ее *функциональная мощность* (множество всех решаемых ею задач, определяющих уровень ее самоорганизации, уровень ее интеллекта), но и уровень ее *обучаемости* (уровень способности к собственной эволюции, к повышению своего уровня самоорганизации, уровня интеллекта). Эволюция *Глобальной экосистемы интероперабельных интеллектуальных компьютерных систем* осуществляется одновременно на следующих архитектурных уровнях этой экосистемы:

- В рамках каждой *индивидуальной интеллектуальной компьютерной системы*, входящей в состав экосистемы (индивидуальная интеллектуальная компьютерная система – это целостная интеллектуальная компьютерная система, способная самостоятельно взаимодействовать с окружающей ее средой, в том числе и с себе подобными);
- В рамках каждого входящего в экосистему *коллектива индивидуальных интеллектуальных компьютерных систем*, входящих в состав экосистемы (каждый такой коллектив представляет собой многоагентную систему, способную коллективно решать соответствующее множество задач);
- В рамках каждого входящего в экосистему *иерархического коллектива интеллектуальных компьютерных систем*, в состав каждого из которых могут входить и индивидуальные интеллектуальные компьютерные системы, и коллективы индивидуальных интеллектуальных компьютерных систем, и другие иерархические коллективы интеллектуальных компьютерных систем;
- В рамках каждой *популяции интеллектуальных компьютерных систем*, входящих в экосистему и автоматически порождаемых и обучаемых в рамках экосистемы.

Рассмотрим подробнее направления эволюции индивидуальных и коллективных (многогенных) кибернетических систем:

- Каждая индивидуальная кибернетическая система (индивидуум, особь, организм) эволюционирует (совершенствуется, самообучается) путем накопления информации об окружающей среде и приобретения опыта по взаимодействию с этой средой. Целью эволюции (самообучения) является выявление таких полезных факторов окружающей среды, которые способствуют сохранению целостности и увеличению срока жизни индивидуальной кибернетической системы, а также выявление опасных (вредных) для кибернетической системы факторов окружающей среды.
- Каждая индивидуальная кибернетическая система имеет конечный срок жизненного

цикла. Поэтому важнейшим этапом эволюции индивидуальных кибернетических систем является приобретение ими способности порождать (воспроизводить) себе подобных (например, путем деления). В результате вместо множества разнообразных индивидуальных кибернетических систем возникает множество разнообразных популяций индивидуальных кибернетических систем, каждая из которых состоит из индивидуальных кибернетических систем одного вида, способных размножаться, т. е. порождать «потомство» новых себе подобных индивидуальных кибернетических систем, которым передается опыт, накопленный «предками».

- Переход от индивидуальных кибернетических систем к их *коллективам*, обеспечивающим существенное расширение множества решаемых задач. В основе таких коллективов лежат принципы организации взаимодействия индивидуальных кибернетических систем при коллективном решении сложных задач.
- Формирование (создание) для всех членов коллектива или популяции кибернетических систем *общей памяти*, в которой аккумулируются все накопленные ими знания и передаются всем членам коллектива или популяции кибернетических систем (в частности, от предков к потомкам). В основе указанной общей памяти лежит общий язык.
- Эволюция кибернетической системы, осуществляемая не только в форме повышения качества своей базы знаний, но и в форме модернизации собственной физической оболочки, а также в форме изменения своей внешней среды.

Основная цель *кибернетики* – познать законы эволюции (самоорганизации), которые в природе реализуются методом проб и ошибок, требуя много времени, и создать искусственные самоорганизующиеся системы, в рамках которых существенно ускорятся темпы их эволюции.

СПИСОК ЛИТЕРАТУРЫ

1. Поспелов, Д. А. Ситуационное управление. Теория и практика / Д. А. Поспелов. – М. : Наука, 1986.
2. Палагин, А. В. Проблемы трансдисциплинарности и роль информатики / А. В. Палагин // Кибернетика и системный анализ. – 2013. – № 5. – С. 3–13.
3. Голенков, В. В. Открытая технология онтологического проектирования, производства и эксплуатации семантически совместимых гибридных интеллектуальных компьютерных систем / В. В. Голенков, Н. А. Гулякина, Д. В. Шункевич. – Минск : Бестпринт, 2021. – 690 с.
4. Тарасов, В. Б. От многоагентных систем к интеллектуальным организациям / В. Б. Тарасов. – М. : Изд-во УРСС, 2002. – 352 с.

RISC-V – ПРОЦЕССОРНЫЕ ТЕХНОЛОГИИ БЕЗ ОГРАНИЧЕНИЙ

Якушкин С. И., Владимиров К. И., Пузикова В. В.

ООО «Синтакор»

Москва, Россия

E-mail: {sergey.yakushkin, konstantin.vladimirov}@syntacore.com

В работе рассматривается открытая процессорная архитектура RISC-V, этапы её развития и стандартизации, ключевые факторы успеха. Обсуждаются понятия системы команд, модульности и расширяемости, профилей и платформ. Проводится анализ мировых трендов и актуальных продуктов на основе архитектуры RISC-V для различных приложений. Кроме того, рассматривается поддержка инструментов разработки и программного обеспечения, а также задачи разработки открытых стандартов для систем с открытой экосистемой приложений и программного обеспечения.

ВВЕДЕНИЕ

Проектирование и разработка современных микропроцессорных ядер – трудоемкая задача, требующая нетривиальных инструментов и серьёзных инвестиций. Сложность разработки новых поколений процессорных архитектур постоянно возрастает из-за увеличивающихся требований производительности, а также по причине того, что большинство распространённых систем команд проприетарны и не расширяемы, т.е. обладают как патентными, так и техническими ограничениями на переиспользование. Это приводит к закреплению различных неудачных решений в экосистеме и сдерживанию развития микроархитектуры и программного обеспечения [1]. Решением перечисленных проблем являются новые принципы построения системы команд и экосистемы разработки, а также переход к свободным и открытым архитектурам [2].

I. ОТКРЫТАЯ ПРОЦЕССОРНАЯ АРХИТЕКТУРА

RISC-V – свободная и открытая процессорная архитектура. Её основой является модульная и расширяемая система команд [3], развиваемая открытым сообществом. Система команд RISC-V спроектирована для широкого спектра оборудования от смарт-карт до высокопроизводительных серверов и ускорителей. Базовый набор команд очень минималистичен, поэтому его легко реализовать в любых решениях. При этом в кодировке оставлено место для стандартных и пользовательских расширений, предназначенных для достижения максимальной эффективности на высокопроизводительных платформах.

Базовый набор команд – это обязательная часть системы команд. Доступно несколько вариантов этого набора. Так, RV64I является основным набором инструкций для работы с 64-битными данными и адресами. Существуют альтернативные варианты базового набора – RV32E, RV32I, RV64E, RV128I. В наборе RV32I присутствует всего лишь 47 инструкций.

Среди основных стандартных расширений можно отметить расширение M для умножения и

деления целых чисел, F и D – для чисел с плавающей точкой одинарной и двойной точности соответственно, V – для векторных команд, B – для битовых операций. Новые стандартизованные расширения помечаются буквами Z, S или H в начале названия. Например, Ztso определяет последовательную консистентность памяти как расширение и т.д. Таким образом, могут существовать RISC-V системы с поддержкой и без поддержки модели памяти TSO (Total Store Ordering).

Дополнительно определяются понятия профилей и платформ. Стандартные профили задают список расширений системы команд обязательных в выбранной области применения. Например, профили RVA20U64 и RVA20S64 для прикладных процессоров объединяют все расширения, которые требуются для загрузки Linux и работы стека приложений на 64-битной архитектуре. Профили обеспечивают гибкость стандарта без фрагментации экосистемы. Стандартные платформы основаны на профилях и дополнительно определяют системные требования программного обеспечения к аппаратному обеспечению, процессу загрузки, управлению питанием и другим аспектам.

II. ЭТАПЫ РАЗВИТИЯ СТАНДАРТОВ И ЭКОСИСТЕМЫ

Проектирование системы команд началось в академической среде приблизительно в 2010 г. Работы включали в себя эксперименты и исследование существующих архитектур, формальное определение RISC-V, создание тестовых SnK. Международный консорциум RISC-V был создан в 2015 г. Уже к 2017 году на рынке появились встроенные ядра для коммуникаций и управления питанием, началось активное развитие программного обеспечения. Основные стандарты были приняты к 2019 году, что позволило создавать микроконтроллерные ядра и SnK для IoT приложений.

На примере линейки микроконтроллерных ядер Syntacore (SCR1, SCR3, SCR4 и SCR6), представленной на рис. 1, можно отметить конкурентные уровни производительности RISC-V, доступ-

ные решения с открытым кодом [4], широкий набор инструментов программирования и библиотек, возможность строить специализированные многопроцессорные системы. Выпущены продукты для широкого диапазона приложений, включающие как микроконтроллеры с низким энергопотреблением, так и высокопроизводительные машины для ЦОД. На китайском рынке можно отметить платежные терминалы Alipay на основе RISC-V, и чип LTE Cat.1 для IoT от China Mobile. По прогнозам Semico Research к 2025 году микроконтроллерные ядра RISC-V займут более 28% мирового рынка IoT.

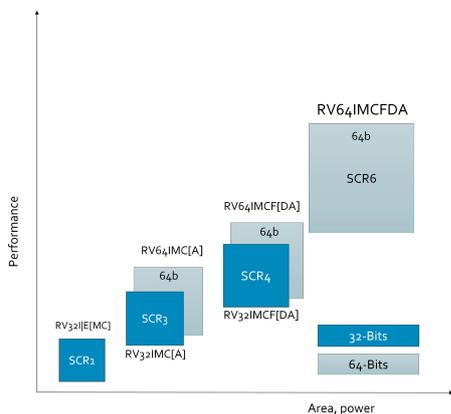


Рис. 1 – Линейка микроконтроллерных ядер SCR

Важной частью развития экосистемы является развитие инструментов разработки. RISC-V поддержан как целевая архитектура в компиляторах GCC и LLVM, в операционных системах Linux и FreeBSD, в системе реального времени Zephyr, в отладчиках GDB и OpenOCD, в большом количестве инструментов разработки. Также доступны модели разного уровня, от функциональных (SAIL, Spike, QEMU) до потактовых, таких как gem5.

III. МОБИЛЬНЫЕ И СЕРВЕРНЫЕ ПРИМЕНЕНИЯ

В 2021 году были приняты стандарты и ратифицированы расширения для многопроцессорных систем с поддержкой Linux, векторных и битовых операций, гипервизора, режима отладки. Отметим, что поддержка Linux имеется и в производимых Syntacore ядрах SCR5, SCR7 и SCR9 (рис. 2), которые используются в системах промышленной автоматизации, медиапроцессорах, а также в процессорах для ИИ.

В 2023 году компания Google объявила о начальной поддержке RISC-V в операционной системе Android. На RISC-V China Summit были представлены серверные процессоры с 64 ядрами SG2042, сервер 8xHS-S1 с 2048 ядрами от PerfXLab, 10 Гбит коммутатор Milk-V Vega, относящиеся к сегменту высокопроизводительных систем [5]. Сообщество активно обсуждает дальнейшее развитие RISC-V: профиль RVA23, серверную и Android платформы, стандарты безопасности

и защиты от атак по сторонним каналам, развитие WorldGuard и IOPMP, разработку Vector 2.0 с 64-битными инструкциями и мн.др.

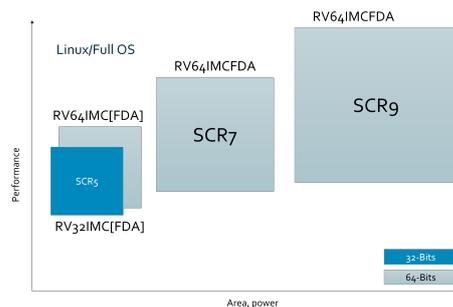


Рис. 2 – Линейка Linux-capable ядер SCR

IV. ЗАКЛЮЧЕНИЕ

Основным фактором успеха RISC-V является открытая кооперация в международном консорциуме, включающем 3500+ участников из 70+ стран. В консорциуме представлено 12 сегментов рынка, 100+ ведущих производителей процессоров (Intel, Qualcomm), 50+ ведущих разработчиков программного обеспечения (Google, Microsoft) и 150+ исследовательских организаций. Сильные региональные сообщества созданы в Китае, Индии, Европе, а с 2022 г. начал свою работу и Российский Альянс RISC-V.

Влияние RISC-V на индустрию можно сравнить с появлением ключевых открытых стандартов и сообществ, таких как USB, PCI и Linux. На данный момент отмечается высокий спрос во всех сегментах и на всех уровнях производительности, производятся существенные инвестиции. Благодаря совместным усилиям большого числа ведущих экспертов и компаний происходит быстрый рост экосистемы и активное развитие технологий. Можно сделать вывод, что успех RISC-V неизбежен не только в сегменте микроконтроллерных ядер, но и в области вычислений общего назначения.

V. СПИСОК ЛИТЕРАТУРЫ

1. Фролов, В. А. Исследование технологии RISC-V / В. А. Фролов, В. А. Галактионов, В. В. Санжаров // Труды Института системного программирования РАН. – 2020. – Т. 32, № 2. – С. 81–98.
2. Программная инфраструктура RISC-V [Электронный ресурс] / Синтакор. – Режим доступа: https://syntacore.com/media/riscv_moscow_2019/riscv_sw_upd.pdf – Дата доступа: 07.11.2023.
3. Design of the RISC-V Instruction Set Architecture: Technical Report / Electrical Engineering and Computer Sciences, University of California at Berkeley; A. Waterman. – Berkeley, 2016. – 104 с. – № UCB/EERC-2016-1.
4. SCR1 RISC-V Core [Electronic resource] / Syntacore. – Mode of access: <https://github.com/syntacore/scr1>. – Date of access: 07.11.2023.
5. RVSC2023 [Electronic resource] / CNRV. – Mode of access: <https://github.com/cnr/vrisc2023>. – Date of access: 07.11.2023.

Секция «Системы управления»

Section «Control systems»

ВЫБОР ОПТИМАЛЬНОГО АЛГОРИТМА ОБРАБОТКИ СИГНАЛОВ ИНЕРЦИАЛЬНОГО ДАТЧИКА ИЗ СОСТАВА РАБОЧЕГО МЕСТА БАЛАНСИРОВКИ РОТОРОВ

Арбузов Г. Н., Сипович А. С., Шобик И. П.

Кафедра систем управления,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {ms.alinasipovich, hleb.arbuzau, innash1820}@gmail.com

В данной работе рассматриваются различные варианты алгоритмов обработки сигнала на стенде балансировки роторов. Описывается алгоритм работы стенда, содержание и итог проведенных экспериментов. Рассматриваются два класса цифровых фильтров: с бесконечной и конечной импульсной характеристикой. В экспериментах используются различные их комбинации. Цель работы: определить наилучший алгоритм обработки данных.

ВВЕДЕНИЕ

Балансировка роторов — операция по уменьшения дисбаланса путем добавления уравновешивающих масс. Задачей балансировки является нахождение величины и места (угла) установки уравновешивающих масс.

Балансировочный стенд включает в себя датчик ускорения, который фиксирует перемещение пластины с закрепленным на ней ротором. Данные передаются на контроллер, который их обрабатывает, фильтрует и выводит на дисплей.

Возникает задача первичной обработки информации в ИИК — фильтрация сигнала $y(t)$ измерительной информации от случайной помехи $e(t)$. В основном, методы фильтрации основаны на различии частотных спектров функций $y(t)$ и $e(t)$, и помеху считают более высокочастотной.

Способ обработки сигналов на основе численных методов с использованием цифровой вычислительной техники — цифровая обработка сигналов. Цифровые фильтры, как и аналоговые, обладают различными характеристиками — частотные: АЧХ, ФЧХ, временная: импульсная характеристика, а также передаточная характеристика фильтра.

Рассматривается два класса фильтров — с бесконечной импульсной характеристикой (БИХ) и конечной импульсной характеристикой (КИХ).

I. ОПИСАНИЕ БАЛАНСИРОВОЧНОГО СДЕНДА

Нижняя пластина столика для балансировки фиксируется на столе при помощи зажима (струбцина). При этом на столе не должно быть приборов, которые могут быть источниками вибраций. Стол должен жестко и устойчиво стоять на полу, чтобы минимально воспринимать колебания.

Верхняя пластина столика крепится к нижней пластине через упругие пластины, таким образом, что она имеет возможность малого линейного перемещения при этом только в одном направлении.

На верхней пластине столика для балансировки жестко закреплен датчик комбинированный инерциальный MPU6050, который имеет в составе акселерометр. Акселерометр воспринимает ускорения колебаний верхней пластины столика по одной оси (X). На верхней пластине жестко закреплена переходная пластина, предназначенная для фиксации балансируемой рамки подвеса.

При вращении рамки, которая жестко связана с ротором двигателя, создаются колебания, которые отражаются на перемещении верхней пластины столика вдоль оси X. Перемещение верхней пластины как амплитуду ускорения фиксирует датчик, который впоследствии передает данные на контроллер Black Pill, который в свою очередь эти данные обрабатывает, фильтрует и выводит на дисплей.

Программа обрабатывает данные с датчика; на основании данных ускорения с акселерометра, изменяющихся по синусоидальному закону, программа рассчитывает частоту вращения двигателя, а также амплитуду ускорения акселерометра и выводит эти данные на дисплей.

II. ТЕОРЕТИЧЕСКОЕ ОПИСАНИЕ ЦИФРОВЫХ ФИЛЬТРОВ

КИХ фильтр (англ. FIR — «finite impulse response») — это цифровой фильтр с конечной импульсной характеристикой. Импульсная характеристика такого фильтра ограничена во времени, то есть имеет счётное число коэффициентов. В общем случае, КИХ-фильтры реализуются без обратных связей, то есть они нерекурсивные.

БИХ фильтр (англ. IIR — «infinite impulse response») — это цифровой фильтр с бесконечной во времени импульсной характеристикой, то есть имеет очень большое или бесконечное число коэффициентов. БИХ фильтры также называют рекурсивными в связи с тем, что при их реализации используются обратные связи.

III. СОДЕРЖАНИЕ ЭКСПЕРИМЕНТА

В ходе эксперимента было реализовано 4 алгоритма обработки данных (рис.1):

1. Сигнал ускорения с акселерометра проходил через БИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего значения сигнала и наименьшего, проходила через БИХ фильтр;
2. Сигнал ускорения с акселерометра проходил через КИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего значения сигнала и наименьшего, проходила через БИХ фильтр;
3. Амплитуда сигнала ускорения с акселерометра, рассчитанная как разница наибольшего значения сигнала и наименьшего, проходила через БИХ фильтр;
4. Сигнал ускорения с акселерометра проходил последовательно через БИХ фильтр и через КИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего значения сигнала и наименьшего, проходила через БИХ фильтр.

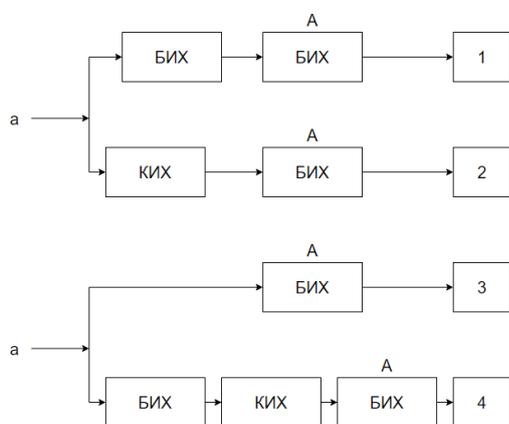


Рис. 1 – Алгоритмы обработки ускорения

Итоги: В ходе эксперимента оценивали отношение сигнал/шум. Выяснили, что наилучшим является 4-ый алгоритм, так как при его использовании получили самое большое значение отношения (рис. 2).

ЗАКЛЮЧЕНИЕ

Рассмотрели различные варианты алгоритмов обработки сигнала на стенде балансировки роторов. В ходе эксперимента выяснили, что наилучшим является алгоритм, в котором сигнал ускорения с акселерометра проходил последовательно через БИХ фильтр и через КИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего значения сигнала и наименьшего, проходила через БИХ фильтр.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Горбачев, А. А., Крюков, А. В., Зуев, В. И. Цифровая обработка сигналов. - Учебное пособие. М.: Высшая школа, 2012.
2. Зуев, В. И., Якубович, И. М., Крюков, В. И. Цифровые фильтры. - Учебное пособие. М.: Высшая школа, 2012.
3. Ковалёв, В. К. Балансировка роторов. - Монография. М.: Машиностроение, 1977.
4. Соболев, И. А., Абрамов, В. В. Балансировка роторов в собственных опорах. - Монография. М.: Машиностроение, 1980.

	Сбалансированный двигатель (А)	Сбаланс.дв. с минимальным грузом (В)			Разность (X = B-A)	$\Delta \cdot 10000$	$X \cdot 10000$	$X \cdot 10000 / \Delta \cdot 10000$
1-ый алгоритм(БИХ(a) + БИХ(A))								
Амплитуда ускорения	0,00285	0,0034			0,00055			
Амплитуда шума Δ	0,00015	0,0004				4	5,5	1,375
2-ой алгоритм(КИХ(a) + БИХ(A))								
Амплитуда ускорения	0,0045	0,006			0,0015			
Амплитуда шума Δ	0,0006	0,0006				6	15	2,5
3-ий алгоритм(БИХ(A))								
Амплитуда ускорения	0,021	0,023			0,002			
Амплитуда шума Δ	0,02	0,01				100	20	0,2
4-ый алгоритм(БИХ(a) + КИХ(a) + БИХ(A))								
Амплитуда ускорения	0,0017	0,0024			0,0007			
Амплитуда шума Δ	0,0002	0,0002				2	7	3,5

Рис. 2 – Данные эксперимента

СРАВНЕНИЕ АЛГОРИТМОВ ОБРАБОТКИ СИГНАЛА АКСЕЛЕРОМЕТРА ИСПОЛЬЗУЕМОГО ДЛЯ ОПРЕДЕЛЕНИЯ НЕУРАВНОВЕШЕННОСТИ РОТОРА В СТЕНДЕ БАЛАНСИРОВКИ

Арбузов Г. Н., Сипович А. С., Шобик И. П.

Кафедра систем управления,

Белорусский государственный университет информатики и радиоэлектроники, Минск, Республика Беларусь

ОАО Пеленг, Минск, Республика Беларусь

E-mail: {ms.alinasipovich, hleb.arbuzau, innash1820}@gmail.com

В данной работе сравниваются алгоритмы обработки сигнала на стенде балансировки роторов. Описывается принцип работы стенда, содержание и итоги проведенных экспериментов. Рассматриваются два цифровых фильтра: с бесконечной и конечной импульсной характеристикой, а также их сочетания. В ходе эксперимента проверяются 4 различных алгоритма обработки. Цель работы: определить алгоритм обработки сигнала, который позволяет производить балансировку с максимальной точностью.

ВВЕДЕНИЕ

Балансировка роторов предназначена для изменения дисбаланса путем добавления уравновешивающих масс. Задачей балансировки является нахождение величины и места установки одной или нескольких уравновешивающих масс.

Ранее для определения неуравновешенности использовался дорогостоящий универсальный измеритель вибраций. В статье рассматривается применение более простого оборудования для снижения стоимости рабочего места. Ставится задача выбора оптимального алгоритма обработки данных.

Сравнивается эффективность двух фильтров – с бесконечной импульсной характеристикой (БИХ) и конечной импульсной характеристикой (КИХ), а также их сочетаний.

I. ОПИСАНИЕ БАЛАНСИРОВОЧНОГО СДЕНДА

Стенд (рис.1) представляет из себя балансировочный столик, состоящий из верхней, нижней и упругих пластин. В состав также входит датчик ускорения, контроллер для обработки данных, драйвер управления двигателем, источник питания.

Верхняя пластина столика крепится к нижней пластине при помощи упругих пластин, таким образом, что она имеет возможность малого линейного перемещения только по одной оси (X).

Нижняя пластина столика для балансировки фиксируется на столе при помощи зажима.

На верхней пластине столика для балансировки жестко закреплен датчик комбинированный инерциальный, который имеет в своем составе акселерометр. Акселерометр воспринимает ускорения колебаний верхней пластины столика по оси X. На верхней пластине жестко закреплен двигатель. При вращении ротора возникают колебания, которые создают колебания верхней пластины столика вдоль оси X. Перемещение верхней пластины как ускорение фиксирует дат-

чик, который впоследствии передает данные на контроллер, который в свою очередь эти данные обрабатывает, фильтрует и выводит на дисплей.

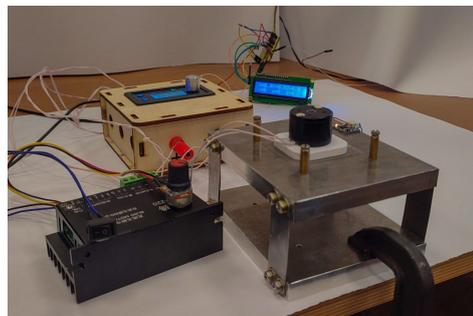


Рис. 1 – Стенд для балансировки

Программное обеспечение контроллера обрабатывает данные с датчика ускорения, рассчитывает амплитуду ускорения и выводит эти данные на дисплей. Данные с акселерометра считываются с частотой 1 кГц. Для акселерометра установлена полоса пропускания 260 Гц. Во время балансировки ротор двигателя вращается с частотой 23 Гц.

II. ОПИСАНИЕ ЦИФРОВЫХ ФИЛЬТРОВ

С целью уменьшения помех, которые возникают при передачи данных с датчика на контроллер, пропустим полученный сигнал через фильтры, описание которых приведено ниже.

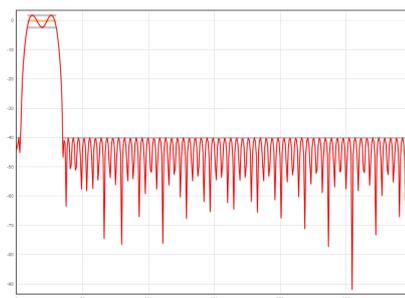


Рис. 2 – АЧХ КИХ-фильтра

Первый фильтр (Ф1), который используется для обработки сигнала акселерометра, – КИХ фильтр, нижняя граница полосы пропускания равна 8 Гц, верхняя граница полосы пропускания – 30 Гц. АЧХ фильтра представлена на рисунке 2.

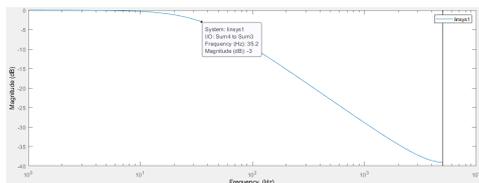


Рис. 3 – ЛАЧХ НЧ БИХ-фильтра (Ф2)

Второй фильтр (Ф2) – низкочастотный БИХ фильтр, полоса пропускания – 35 Гц. ЛАЧХ фильтра представлена на рисунке 3.

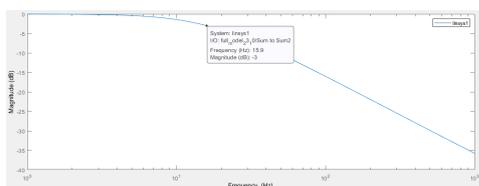


Рис. 4 – ЛАЧХ НЧ БИХ-фильтра (Ф3)

Третий фильтр (Ф3) – низкочастотный БИХ фильтр, полоса пропускания – 16 Гц. ЛАЧХ фильтра представлена на рисунке 4.

III. СОДЕРЖАНИЕ ЭКСПЕРИМЕНТА

Эксперимент включает в себя две части: измерение амплитуды ускорения при сбалансированном роторе и измерение амплитуды ускорения сбалансированного ротора с минимальным грузом.

В ходе эксперимента было проверено 4 алгоритма обработки сигнала (рис.5):

1. Амплитуда сигнала ускорения с акселерометра, рассчитанная как разница наибольшего и наименьшего значений сигнала, проходила через БИХ фильтр;
2. Сигнал ускорения с акселерометра проходил через КИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего и наименьшего значений сигнала, проходила через БИХ фильтр;

3. Сигнал ускорения с акселерометра проходил через БИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего и наименьшего значений сигнала, проходила через БИХ фильтр;
4. Сигнал ускорения с акселерометра проходил последовательно через БИХ фильтр и через КИХ фильтр, а амплитуда отфильтрованного сигнала, рассчитанная как разница наибольшего и наименьшего значений сигнала, проходила через БИХ фильтр.

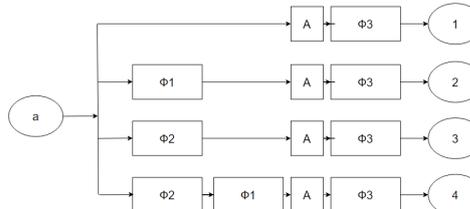


Рис. 5 – Алгоритмы обработки сигнала

Итоги: в ходе эксперимента оценивали отношение сигнал/шум. Выяснили, что наилучшим является 4-ый алгоритм, так как при его использовании получили самое большое значение отношения сигнал/шум (таблица 1).

IV. ЗАКЛЮЧЕНИЕ

Рассмотрены различные варианты алгоритмов обработки сигнала на стенде балансировки роторов. В ходе эксперимента выбран оптимальный алгоритм обработки данных, который позволяет снизить шумы исходного датчика в 20 раз, что дает возможность определить неуравновешенность ротора с точностью 0,5 г*мм. Применение данного стенда позволило снизить стоимость рабочего места для балансировки в десятки раз.

СПИСОК ЛИТЕРАТУРЫ

1. Зуев, В. И., Якубович, И. М., Крюков, В. И. Цифровые фильтры. – Учебное пособие. М.: Высшая школа, 2012.
2. Ковалёв, В. К. Балансировка роторов. – Монография. М.: Машиностроение, 1977.
3. Соболев, И. А., Абрамов, В. В. Балансировка роторов в собственных опорах. – Монография. М.: Машиностроение, 1980.
4. ГОСТ 19534-74. Балансировка вращающихся тел.

Таблица 1 – Данные эксперимента

№ алг.	Изм. параметры	Без груза	С грузом	Разность	Сигнал/ шум
1	Ампл., mg	21,00	23,00	2,00	
	Шум, mg	1,00	1,00		2,00
2	Ампл., mg	4,50	6,00	0,15	
	Шум, mg	0,60	0,60		2,5
3	Ампл., mg	2,85	3,40	0,55	
	Шум, mg	0,20	0,20		2,75
4	Ампл., mg	1,70	2,40	0,70	
	Шум, mg	0,20	0,20		3,50

ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ В СОВРЕМЕННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ БЕСПИЛОТНЫМИ АВТОМОБИЛЯМИ

Ахапкина А. М., Способ С. П.

Кафедра связи, Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {anastasia.akhapkina2018}@gmail.com

В данной статье описаны довольно широкое применение нейросетевых технологий в современных системах управления беспилотным автомобилем, задачей которого является безопасное транспортное передвижение на дорогах общего пользования, избегая различных препятствий, а также соблюдение правил дорожного движения. В статье обосновывается актуальность и практичность использования нейронных сетей в данной системе управления.

ВВЕДЕНИЕ

Современный мир – это эра информационных технологий и высокой доступности информации. В последние два десятилетия наблюдается стремительное развитие информационных структур и их интеграция в повседневную жизнь человека. Сегодня невозможно представить себе обыденный день без использования информационных технологий, будь то в виде виртуальных ассистентов, смарт-домов или автономных автомобилей. Эти передовые технологии, называемые искусственным интеллектом (ИИ), значительно упрощают выполнение разнообразных периодических задач, которые ранее требовали сложных алгоритмических решений и большого количества времени [1].

Один из наиболее перспективных методов искусственного интеллекта, который заменяет труд человека и обеспечивает обработку информации схожей с работой человеческого мозга, – это использование нейронных сетей. Нейронные сети представляют собой математическую модель, а также программное или аппаратное воплощение этой модели, основанную на принципах организации и функционирования биологических нейронных сетей, которые присущи живым организмам [2]. В большинстве случаев нейронные сети реализуются в виде компьютерных программ или технологических решений.

I. ОСНОВНАЯ ЧАСТЬ

Одним из важнейших аспектов управления беспилотными транспортными средствами является необходимость точно знать окружающую среду и препятствия, которые могут возникнуть на пути. Эти препятствия могут представлять собой движущиеся автомобили, здания, пешеходов на пешеходных переходах и различные объекты. Для обеспечения безопасного движения беспилотных транспортных средств разрабатываются алгоритмы, основанные на системах автономного вождения. Одним из ключевых элементов в этих алгоритмах являются сверточные нейронные се-

ти, которые специализируются на распознавании объектов на дороге в реальном времени [3].

Для обеспечения правильной навигации беспилотного автомобиля также необходимо знать расстояние до заранее запрограммированной цели. Эта информация является фундаментальной для корректной работы нейронной сети. Беспилотные автомобили оснащены разнообразными датчиками и сенсорами, которые непрерывно собирают данные и передают их в нейронную сеть. Эти данные включают в себя информацию о расположении препятствий в пространстве. Нейронная сеть в свою очередь определяет не только направление движения, но и соответствующие ускорения или замедления.

Процесс обучения нейронной сети играет важную роль в обеспечении верного и безопасного движения автомобиля. Нейронная сеть обучается на основе входных и выходных данных. Входы представляют собой информацию о расположении препятствий, а выходы определяют управление движением автомобиля.

Для обеспечения безопасного и эффективно автономного движения автомобиля необходимо начать с обучения и настройки искусственных нейронных сетей. Электронные средства транспорта оборудованы комплексными информационными системами, включающими множество датчиков и сенсоров, которые непрерывно взаимодействуют с нейронной сетью и предоставляют данные для анализа.

Исходя из этой информации, можно определить параметры и структуру нейронной сети, включая количество входов и соответствующих им выходов. Входные данные могут включать информацию о местоположении препятствий вокруг автомобиля, в то время как выходные данные могут управлять направлением движения и скоростью.

Успешность функционирования автономного автомобиля зависит от того, как эффективно обучена нейронная сеть. Неправильное обучение может привести к нежелательным действиям или даже авариям. Для гарантированной безопас-

ности и надежности, обучение нейронных сетей должно быть внимательно спланировано и настроено. Примерные параметры обучения нейронной сети могут быть найдены в соответствующей таблице (см. рисунок 1) [4].

Входные сигналы			Выходные сигналы	
Относительное расстояние до препятствия				
По левую сторону	По центру	По правую сторону	Ускорение	Направление
Нет препятствия	Нет препятствия	Нет препятствия	Ускорение вперед	Прямо
Четверть пути	Нет препятствия	Объект близко	Не большой разгон	Чуть левее
Задевание объекта	Объект близко	Задевание объекта	Движение назад	Налево

Рис. 1 – Матрица расположения препятствий относительно автомобиля и ответ нейронной сети

Другим эффективным подходом к управлению автономным автомобилем является использование контроллера на основе нейронных сетей, который способен адаптироваться к различным внешним воздействиям, таким как окружающие препятствия. Этот контроллер принимает входные сигналы о состоянии системы и воздействиях на нее, а затем выдает соответствующие управляющие сигналы для механических устройств (см. рисунок 2) [5].

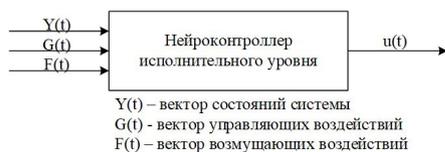


Рис. 2 – Схема нейроконтроллера

II. ЗАКЛЮЧЕНИЕ

Использование нейронных сетей в автономных автомобилях предоставляет ряд преимуществ и недостатков. С одной стороны, автома-

тизация управления автомобилем может значительно улучшить качество жизни людей, освобождая их от монотонной работы водителей такси и грузоперевозчиков. Правильно настроенные нейросетевые технологии могут снизить количество дорожно-транспортных происшествий и, как следствие, спасти жизни. Кроме того, это может привести к экономии времени и снижению затрат на оплату труда водителей, перевозящих пассажиров и грузы, а также к сокращению рабочих мест, связанных с водительской профессией. Следует отметить, что автономные автомобили, управляемые искусственными нейронными сетями, могут иметь более высокую стоимость из-за сложности электронных систем и датчиков, необходимых для обеспечения безопасности и надежности. Кроме того, нейронные сети не всегда способны адекватно реагировать на критические ситуации, которые могут возникнуть на дорогах. Поэтому разработчики должны продолжать работу над совершенствованием этой технологии, чтобы обеспечить безопасность и надежность автономных автомобилей в самых разнообразных условиях.

СПИСОК ЛИТЕРАТУРЫ

1. Цаунит А.Н. Перспективы развития и применения нейронных сетей / А.Н. Цаунит. – Текст: непосредственный // Молодой ученый. – 2021. – №23 (365).
2.] Галушкин А.И. Нейронные сети: основы теории. / А.И. Галушкин. М.: РиС, 2015. 496 с.
3. Сверточная нейронная сеть [Электронный ресурс]: Википедия. Свободная энциклопедия.
4. M.T. Hagan, H.V. Demuth and M.H. Beale, Neural Network Design, PWS Publishing, Boston, MA, 1995
5. Осипов Г.С. Оптимизация одноканальных систем массового обслуживания с неограниченной очередью// Бюллетень науки и практики. 2016. №9(10)

НЕЙРОННЫЕ СЕТИ В АВТОНОМНОМ УПРАВЛЕНИИ: ПУТЬ К БЕЗОПАСНЫМ И ЭФФЕКТИВНЫМ ТРАНСПОРТНЫМ СРЕДСТВАМ

Ахапкина А. М., Стежко И. К.

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {anastasia.akhapkina2018}@gmail.com

В данной статье рассматривается физическая и программная реализация автономного управления с использованием нейронных сетей, а также предоставим реальные примеры автомобилей, где эта система успешно применяется.

ВВЕДЕНИЕ

Автономное управление автомобилями является фундаментальной областью исследований и разработки в сфере транспортных технологий. Нейронные сети и глубокое обучение играют ключевую роль в реализации этой концепции.

Зачем нужно автономное управление? Автономное управление представляет собой ответ на существующие вызовы и потребности в области транспорта и транспортной безопасности. Первой важной причиной его разработки является улучшение безопасности дорожного движения. Автономные системы могут обнаруживать и реагировать на опасные ситуации быстрее и точнее, чем человек, что снижает количество аварий и спасает жизни. Второй важной причиной – увеличение эффективности дорожного движения. Автономные автомобили могут двигаться ближе друг к другу, оптимизируя поток транспорта и уменьшая пробки, что также ведет к снижению выбросов и более эффективному использованию инфраструктуры.

I. ПРЕИМУЩЕСТВО НЕЙРОННЫХ СЕТЕЙ

Нейронные сети стали ключевым элементом в разработке автономных систем. Их способность обучения на больших объемах данных, адаптации к изменяющимся условиям и точности в распознавании окружающей среды делают их идеальными для этой задачи. Нейронные сети могут анализировать данные с множества сенсоров, определять объекты и дорожные условия, а также предсказывать оптимальные решения для безопасного и эффективного движения [1].

II. РОСТ ПОПУЛЯРНОСТИ

Автономное управление с нейронными сетями становится все более популярным по нескольким причинам:

- **Безопасность:** в свете статистики дорожных аварий, автономное управление представляется как способ снизить количество несчастных случаев и спасти жизни[2].
- **Эффективность:** автономные автомобили могут экономить время и топливо благо-

даря оптимизации движения и управлению эффективным маршрутом.

- **Комфорт:** пассажиры могут сосредоточиться на других делах, не беспокоясь о вождении.
- **Экологическая устойчивость:** уменьшение пробок и более эффективное движение ведет к снижению выбросов и уменьшению воздействия на окружающую среду.

III. ТЕХНОЛОГИЧЕСКИЕ ВОЗМОЖНОСТИ

Автономное управление, в сочетании с нейронными сетями, предоставляет ряд впечатляющих технологических возможностей [3]:

- Автоматическое восприятие окружающей среды: Нейронные сети способны анализировать данные с разнообразных сенсоров, включая камеры, радары и лидары, для точного восприятия окружающей среды и обнаружения объектов на дороге.
- Обучение на больших объемах данных: глубокое обучение нейронных сетей требует огромных объемов данных, что позволяет им распознавать паттерны и улучшать свои – способности с течением времени.
- Предсказание и планирование: нейронные сети могут предсказывать дорожную обстановку и принимать решения о маневрах, учитывая безопасность и эффективность движения[4].

IV. ТЕХНОЛОГИЧЕСКИЕ ВЫЗОВЫ

С развитием автономного управления также существуют и технологические вызовы:

- **Безопасность:** одним из главных вызовов является обеспечение безопасности. Нейронные сети должны быть надежными и устойчивыми к атакам или сбоям;
- **Нормативное регулирование:** необходимы ясные нормативные стандарты для автономных автомобилей, чтобы обеспечить их безопасное внедрение на дорогах;
- **Этические вопросы:** автономные системы должны принимать этические решения, например, в случаях, когда столкновение неизбежно, выбирая наименее опасное решение [5].

V. ФИЗИЧЕСКАЯ РЕАЛИЗАЦИЯ

Физическая реализация автономного управления начинается с оборудования автомобиля сенсорами. Эти сенсоры включают камеры, радары, лидары и ультразвуковые сенсоры, физически размещенные на автомобиле. Например, Tesla Model S оснащена 8 камерами, 12 ультразвуковыми сенсорами и радаром, расположенными по всей поверхности автомобиля. Физическая интеграция сенсоров требует точной калибровки, чтобы обеспечить точное восприятие окружающей среды.

Следующим шагом физической реализации является актуаторы, которые позволяют автомобилю реагировать на решения, принимаемые программным обеспечением. Это включает двигатели для управления скоростью и рулевое управление для контроля направления движения. Программное обеспечение автомобиля связывает данные с сенсоров с актуаторами, чтобы осуществить физическое управление движением [6].

VI. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Программная реализация автономного управления включает в себя несколько ключевых компонентов. В центре системы находятся нейронные сети, способные анализировать данные с сенсоров и принимать решения. Эти сети обучаются на огромных объемах данных, включая изображения дорожных сцен, данные о движении других участников и информацию о погодных условиях. Обучение нейронных сетей требует вычислительных ресурсов и специальных алгоритмов, которые позволяют им распознавать объекты и предсказывать дорожную обстановку [6].

Дополнительно, алгоритмы машинного обучения применяются для обработки данных с сенсоров и управления актуаторами. Эти алгоритмы позволяют автомобилю реагировать на окружающую среду, соблюдать правила дорожного движения и учитывать взаимодействие с другими участниками движения.

VII. ПРИМЕРЫ РЕАЛЬНЫХ АВТОМОБИЛЕЙ

Применение нейронных сетей в автономных автомобилях уже имеет конкретные реализации. Компания Tesla с ее системой Autopilot представляет реальный пример успешной интеграции нейронных сетей и автономного управления.

Autopilot позволяет автомобилю выполнять функции автоматической езды, включая управление скоростью, удержание полосы движения и выполнение маневров [7].

Waymo (проект Google Self-Driving Car) активно разрабатывает автономные автомобили с использованием нейронных сетей. Эти автомобили успешно тестируются в реальных городских условиях и демонстрируют способность автономного управления.

General Motors совместно с Cruise Automation разрабатывают автономные автомобили, в которых нейронные сети играют важную роль. Эти проекты являются лишь началом внедрения технологии автономного управления на дорогах.

VIII. ЗАКЛЮЧЕНИЕ

Физическая и программная реализация автономного управления автомобилями с использованием нейронных сетей представляет собой сложный и многозадачный процесс. Сенсоры, актуаторы и алгоритмы машинного обучения интегрируются в комплексную систему, способную воспринимать окружающую среду и принимать решения. Примеры реальных автомобилей, такие как Tesla Autopilot и Waymo, демонстрируют, что эта технология уже применяется на практике и представляет собой важный шаг в будущее транспорта. С развитием технологии и нормативного регулирования, автономное управление обещает революционизировать транспортную инфраструктуру, повысить безопасность и улучшить эффективность дорожного движения [2,4].

СПИСОК ЛИТЕРАТУРЫ

1. Xu, Jing, et al. "Deep Learning for Autonomous Driving: A Review"// IEEE Xplore, 2017.
2. NVIDIA. "End-to-End Deep Learning for Self-Driving Cars"// NVIDIA Developer.
3. Shladover, Steven E. "A Survey of Autonomous Vehicles: Challenges, Methods, and Opportunities"// IEEE Xplore, страницы 2448-2457, 2015.
4. Pan, Xingang, et al. "Deep Reinforcement Learning for Autonomous Driving"// arXiv, 2016.
5. Li, Keqiang, et al. "Challenges of Connected and Autonomous Vehicle Technologies"// Amazon, 2019.
6. Ведерников, Александр М. и Давид Ф. Дюфф. "Autonomous Vehicles: Technology, Regulatory, and Policy Issues"// Amazon, 2021.
7. Андрей Карпати и Себастьян Труньен. "Deep Learning for Self-Driving Cars"// Coursera, курс на Coursera, Stanford University.

ПРОБЛЕМЫ АВТОМАТИЗАЦИИ ПРОЦЕССОВ ПРОИЗВОДСТВА МОЛОЧНОЙ ПРОДУКЦИИ

Гурская Я. С., Марков А. В.

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

E-mail: arkipenko.yana.serg@gmail.com., markov@bsuir.by.

В докладе рассмотрены технологические процессы на определенных этапах производства молочной продукции. Приведены общие требования, предъявляемые к процессам упаковки и маркировки готовой продукции. Выявлены проблемы автоматизации перечисленных этапов производства. Предложены варианты решения проблем.

ВВЕДЕНИЕ

Производство молочных продуктов состоит из множества этапов. В зависимости от типа выпускаемой продукции (молоко, сметана, творог, сыр, мороженое и др.) рецептура приготовления различается, что влияет на сложность, общую продолжительность технологического процесса и оснащение производственных линий.

Одним из методов достижения высокой производительности является уменьшение доли ручного труда и автоматизация отдельных процессов на производстве.

Финальным этапом любого производства является упаковка готового изделия и подготовка продукции к транспортировке. Большинство групп молочных продуктов попадают под требования об обязательной маркировке товаров унифицированными контрольными (идентификационными) знаками и средствами идентификации [1, 2, 3]. В связи с данными требованиями, процесс упаковки готовой продукции необходимо модернизировать – предусмотреть установку специального оборудования, которое позволит выполнять маркировку упакованных продуктов в автоматическом режиме.

I. ТРЕБОВАНИЯ К СИСТЕМЕ АВТОМАТИЗАЦИИ

Автоматизация производственных процессов упаковки и маркировки молочной продукции заключается в обеспечении согласованной работы всех узлов и механизмов на этапах упаковки-маркировки, сохранении непрерывного последовательного выполнения указанных процессов, а также соблюдении следующих требований:

- установка нового оборудования без нарушения работы старых системы;
- сохранение скорости работы упаковочных машин на прежнем уровне;
- качественное нанесение унифицированного контрольного знака на единицу продукции;
- считывание кодов маркировки с минимальной потерей количества и качества кадров;
- уменьшение доли ручного труда при настройке и обслуживании нового оборудования;

– возможность работать с товарами различных характеристик (габариты, цвет и материал упаковки, вес и др.)

II. ОБЩИЕ ПРОБЛЕМЫ АВТОМАТИЗАЦИИ

Основная проблема автоматизации процессов производства заключается в обеспечении согласованной работы новых систем и старых, давно успешно работающих на заводах.

Марка и наименование производителя старого оборудования, как правило, не учитываются при проектировании новой системы. Следовательно, очередной проблемой является корректное обслуживание каждого узла системы и обеспечение наличия запасных частей для оборудования различного производства.

Принцип действия оборудования одного типа, но разного производителя аналогичен, однако на практике сталкиваются с проблемой неодинаковой настройки параметров. При обслуживании определенной рабочей системы долгое время, алгоритмы и подходы к настройке оборудования запоминаются человеком и в будущем при внедрении новой системы может возникнуть несоответствие ожидаемых методов настройки оборудования и реальных, а значит, возникает проблема в переобучении персонала.

III. ПРОБЛЕМЫ ЭТАПА МАРКИРОВКИ

Непосредственно на этапе маркировки можно выделить следующие проблемы:

– Работа с продукцией различных характеристик.

Цвет упаковки влияет на качество фото, выполняемое камерой технического зрения при считывании кодов. Необходимо настроить параметры камеры таким образом, чтобы коды считывались корректно при светлой, темной и цветной упаковке.

Габариты продукции влияют на частоту появления кодов маркировки в кадре. Товары больших размеров реже появляются в поле зрения камеры, следовательно, камера в единицу времени выполняет меньше фото, передает для обработки меньше данных, что снижает нагрузку на сеть.

Место нанесения кодов маркировки может различаться в зависимости от типа и габаритов продукции, что необходимо учитывать при размещении и настройке камеры технического зрения. Необходимо добиться настройки параметров облучения таким образом, чтобы система работала корректно для продукции различных типов.

– Ограниченное пространство в производственных помещениях.

Зачастую линии по производству продуктов питания размещаются в давно построенных зданиях. Для размещения минимального количества обязательного оборудования свободного места достаточно. Со временем к процессам производства предъявляются новые требования, что влечет за собой дооснащение существующих систем.

Установка нового оборудования происходит в условиях нехватки свободного места, поэтому приходится разрабатывать комплексы с учетом минимизации занятого пространства, либо при возможности, максимально задействовать поверхности существующего оборудования для закрепления некоторых объектов.

– Оборудование для пищевого производства подвергается санитарной обработке, следовательно, необходимо обеспечить повышенный класс защиты элементов системы.

– В процесс маркировки продукции происходит обработка больших объемов данных. Камеры технического зрения считывают множество единичных кодов маркировки, передают по сети данные контроллерам, которые обрабатывают информацию и передают её на следующий уровень, где данные преобразуются и отображаются пользователю в понятном виде.

Удаленный сервер для сбора данных следует размещать на минимальном расстоянии от комплекса маркировки, чтобы минимизировать сетевые потери. Либо использовать локальный сервер, что позволит минимизировать внешние помехи при передаче данных.

ЗАКЛЮЧЕНИЕ

Несмотря на некоторые проблемы, возникающие в процессе модернизации определенных этапов производства, предприятия совершенствуют линии и автоматизируют технологические процессы, успешно решая поставленные задачи.

Грамотное проектирование новой системы, проверка теоретических разработок на практике и тщательное изучение технической документации, позволяет решить проблему согласования старых и новых узлов на производстве.

Целью разработки любой системы является создание независимого программно-аппаратного комплекса, отвечающего за выполнение процессов на определенном этапе производства, который можно беспрепятственно установить на любом заводе. Наименование производителя и особенности настройки уже имеющегося оборудования не должны оказывать влияние на новый комплекс. При разработке системы предпочтение отдается оборудованию, которое имеет простой, понятный визуальный интерфейс, доступный для освоения рабочим персоналом.

Проблемы маркировки решаются индивидуально в зависимости от предприятия. Общие рекомендации по решению наиболее часто встречающихся сложностей, приведены далее.

– Для улучшения качества фотографии камерой технического зрения и минимизации бликов, рекомендуется задействовать максимальную долю естественного освещения. При невозможности обеспечить поступление естественного света, следует установить специальную подсветку вблизи камеры, а также использовать поляризационные фильтры.

– Проблема настройки системы для работы с продукцией разных габаритов решается программным методом. Следует предусмотреть автоматический выбор типа продукции перед началом работы. В зависимости от выбора определенного варианта, будут установлены различные значения скорости движения конвейеров, частоты фотографирования камерой технического зрения и другие необходимые настройки.

– Повышенный класс защиты элементов системы обеспечивается установкой дополнительных кожухов и защитных экранов для камер, выбором более прочных материалов для изготовления конвейеров, использованием электрооборудования с высокими эксплуатационными характеристиками.

1. О маркировке товаров. Указ Президента Республики Беларусь от 10 июня 2011 г. № 243 // Национальный реестр правовых актов Республики Беларусь, 17.09.2011
2. Указ Президента Республики Беларусь от 6 января 2021 г. № 9 // Национальный правовой Интернет-портал Республики Беларусь, 07.01.2021, 1/19446.
3. О подлежащих маркировке товарах. Постановление Совета Министров Республики Беларусь от 29 июля 2011 г. № 1030 // Национальный реестр правовых актов Республики Беларусь, 02.08.2011 г., № 88, 5/34243

СОВРЕМЕННЫЕ СИСТЕМЫ ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ СООБЩЕНИЙ ЭЛЕКТРОННОЙ ПОЧТЫ

Ломонос Г. В., Захарьев В. А.

Кафедра информационных систем и технологий,
Белорусский государственный технологический университет

Кафедра систем управления,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: lomonosg07@gmail.com, zahariev@bsuir.by

Статья посвящена исследованию современных систем и методов обнаружения вредоносных сообщений в электронной почте. В ней представлен обзор наиболее эффективных подходов и технологий, а также рассмотрены ключевые особенности их реализации.

ВВЕДЕНИЕ

В современном мире, где интернет стал неотъемлемой частью повседневной жизни, компьютерные технологии служат ключевым инструментом доступа к информации. Эти технологии обеспечивают обмен данными, включая электронную почту и мгновенные сообщения. Согласно статистике, к 2023 году число пользователей интернета достигло 4,8 миллиарда, что привело к увеличению нагрузки на сетевые инфраструктуры и к росту объема отправляемой корреспонденции [1]. Однако с ростом использования этих средств коммуникации возникают различные проблемы. Одной из наиболее актуальных является проблема фильтрации спам-писем. Спам-письма можно определить как нежелательные сообщения, которые доставляются на электронные адреса без предварительного согласия получателя. В большинстве случаев это рекламные материалы, распространяемые незаконными методами массовой рассылки [1].

1. СОВРЕМЕННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ СООБЩЕНИЙ

Существует программное обеспечение (ПО) для автоматического определения спама (т. н. спам-фильтры). Оно может быть предназначено для конечных пользователей или для использования на серверах. Программы автоматической фильтрации используют статистический анализ содержания письма для принятия решения, является ли оно спамом. Наибольшего успеха удалось достичь с помощью алгоритмов, основанных на теореме Байеса. Для работы этих методов требуется предварительное «обучение» фильтров путем передачи ему рассортированных вручную писем для выявления статистических особенностей нормальных писем и спама.

Метод очень хорошо работает при сортировке текстовых сообщений (в том числе HTML). После обучения на достаточно большой выборке удаётся отсеять до 95–97 процентов спама. Для обхода таких фильтров спамеры иногда помещают содержательную часть в картинку, вложенную в

письмо, текст же либо отсутствует, либо случаен, что не позволяет фильтру составить статистику для распознавания таких писем. В этом случае необходимо пользоваться программами распознавания текста (большинство современных почтовых программ этого не поддерживают), либо использовать другие методы [2]. Залог надежной работы байесовского метода – постоянное дообучение фильтра и указание ему на совершаемые ошибки. В почтовых программах для этого вводится возможность ручной пометки сообщения «спам/не-спам», а в почтовых сервисах в интернете – кнопка «пожаловаться на спам».

Многие программы и почтовые сервисы в интернете позволяют пользователю задавать собственные фильтры. Такие фильтры могут состоять из слов или, реже, регулярных выражений, в зависимости от наличия или отсутствия которых сообщение попадает или не попадает в мусорный ящик. Однако такая фильтрация трудоёмкая и негибкая, кроме того, требует от пользователя известной степени знакомства с компьютерами. С другой стороны, она позволяет эффективно отсеять часть спама, и пользователь точно знает, какие сообщения будут отсеяны и почему.

Например, одним из методов неавтоматической фильтрации являются черные списки, которые как метод широко используются и поэтому будут кратко освещены, однако не относятся к методам классификации текстовой информации.

Чёрные списки. DNSBL – DNS blacklist или DNS blocklist – списки хостов, хранимые с использованием системы архитектуры DNS. Обычно используются для борьбы со спамом. Почтовый сервер обращается к DNSBL, и проверяет в нём наличие IP-адреса клиента, с которого он принимает сообщение. При положительном ответе считается, что происходит попытка приёма спам-сообщения. Серверу отправителя сообщается ошибка 5xx (неустраняемая ошибка) и сообщение не принимается. Почтовый сервер отправителя создаёт «отказную квитанцию» отправителю о доставке почты. Раньше такие списки назывались RBL, Real-time Blackhole List, но сейчас это название яв-

ляется торговой маркой, принадлежащей MAPS LLC [3].

Типы DNSBL

- Списки открытых релейов – база данных почтовых серверов, неправильно сконфигурированных, которые позволяют пересылать через себя почтовые сообщения для всех желающих. Как правило данные хосты автоматически сканируются в Интернете, поэтому попадание такого хоста в руки людей, рассылающих спам сообщения, происходит очень быстро (не более 4 дней). При использовании данных списков существует наименьшая опасность блокирования обычной почты, так как сервер попадает в список, только после проверки его специальным почтовым роботом.
- Списки спам серверов – база данных серверов, через которые было замечено прохождение спам сообщений. Данные списки составляются на основе показаний пользователей, получивших спам с какого-либо сервера, поэтому они могут содержать устаревшую, или просто неверную информацию [4].
- Список Dialup адресов – список IP адресов провайдеров, используемых ими для организации сервиса удалённого доступа, и, следовательно, которые не могут быть адресами почтовых серверов. Использование данных списков практически безопасно для легальной почты.
- Список открытых HTTP/Socks прокси-серверов без контроля доступа позволяющие любому пользователю совершать неавторизованные действия скрывая свой реальный IP адрес, незаконные действия включают не только рассылку спама, но также и многочисленные иные варианты.

II. СОВРЕМЕННЫЕ СИСТЕМЫ ОБНАРУЖЕНИЯ ВРЕДНОСНЫХ СООБЩЕНИЙ

Одним из самых популярных, универсальных и проверенных средств спам-фильтрации является продукт компании Лаборатории Касперского. Kaspersky Anti-Spam – это решение для защиты пользователей корпоративных почтовых систем и публичных почтовых сервисов от массовой незапрошенной корреспонденции – спама.

В данном программном обеспечении применяются следующие алгоритмы распознавания спама. Проверка сообщения по спискам. Приложение проверяет IP-адрес отправителя по черным спискам провайдеров и общественных организаций (DNSBL – DNS-based Blackhole List). В случае если адрес занесен администратором в белый список, то сообщение принимается, минуя все этапы анализа [5]. Фильтрация по SPF и SURBL. В процессе фильтрации может учитываться авторизация отправителя по технологии

SPF (Sender Policy Framework). В дополнение к спискам DNSBL, выявляющим спамерские IP-адреса, используется также технология SURBL (Spam URI Realtime Block List), распознающая спамерские URL в теле сообщения.

Анализ формальных признаков письма. Программа отсеивает спам по таким типичным для него признакам, как модификация адреса отправителя или отсутствие его IP-адреса в системе доменных имён (DNS), неоправданно большое количество получателей или сокрытие их адресов. Кроме того, оцениваются размер и формат сообщения.

Сигнатурный анализ. Использование круглосуточно обновляемой базы лексических сигнатур позволяет распознавать модифицированные варианты исходного спамерского письма, создаваемые для обхода спам-фильтров.

Лингвистические эвристики. Программа проверяет наличие и расположение в тексте письма слов и фраз, типичных для спама. Анализу подвергается как текст самого письма, так и содержание вложенных файлов. Графические сигнатуры. Используя базу графических сигнатур, приложение блокирует также распространенные в последнее время спамерские письма, которые содержат информацию в виде изображений, а не в виде текста.

UDS-запросы в режиме реального времени. Технология UDS (Urgent Detection System) позволяет получать данные о последних спамерских рассылках уже через секунду после их обнаружения. Эта информация используется для дополнительной проверки тех сообщений, которые не получили однозначной оценки (спам/не-спам) [6].

III. ВЫВОДЫ

В области фильтрации спама используются различные методы распознавания. Однако ни один алгоритм не гарантирует 0 процентов ложноположительных или ложноотрицательных результатов. По мере того, как спамеры совершенствуют свои стратегии рассылки спама, существующие решения по фильтрации спама также должны совершенствоваться.

1. CORMACK, G. V. Email spam filtering: a systematic review. *Foundations and Trends® in Information Retrieval*, 2006, vol. 1, no. 4, pp. 335–455. <https://doi.org/10.1561/1500000006>
2. Global Security Map by SAINT [Электронный ресурс]. – Режим доступа: <<https://globalsecuritymap.com>>
3. Li, Wenbin, Ning Zhong, and Chunnian Liu. "Combining multiple email filters based on multivariate statistical analysis." *Foundations of Intelligent Systems*. Springer Berlin Heidelberg, 2006. 729-738.
4. D.J. Hill, B. S. Minsker, and E. Amir, "Real-time Bayesian anomaly detection in streaming environmental data", *Water Resour. Res.*, 2018, C.45.
5. A. A. Nasr, M. Z. Abdulmaged, "A Learnable Anomaly Detection System using Attributional Rules", *International Journal of Computer Network and Information Security*, vol. 8(11), 2016, C. 57.
6. M. Zhang, B. Xu and J. Gong, "An Anomaly Detection Model Based on One-Class SVM to Detect Network Intrusions," 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN), Shenzhen, 2015, C. 157.

OLAP-ТЕХНОЛОГИИ И ВИЗУАЛИЗАЦИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Маркова А. А., Кузнецов А. П.

Кафедра систем управления,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: 17anamarkov@gmail.com, kuznap@bsuir.by

OLAP (англ. online analytical processing, интерактивная аналитическая обработка) — технология обработки данных, заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу. Реализации технологии OLAP являются компонентами программных решений класса Business Intelligence. Причина использования OLAP для обработки запросов — скорость. Внедрение и развитие технологий бизнес – аналитики на основе концепции визуализации информационных ресурсов с помощью BI-решений позволит эффективно использовать не только финансовые, но человеческие и материальные ресурсы компании.

ВВЕДЕНИЕ

Требования к информационным системам и технологиям при управлении рисками на предприятии включают в себя автоматизированный доступ, консолидацию данных, анализ, прогнозирование и оценку рисков, предоставление полученных результатов. Учитывая многообразие и большой объем данных для анализа, информационные системы и основанные на них системы поддержки, принятия решений являются в настоящее время важнейшим условием эффективного управления рисками в масштабе всей компании.

I. СОВРЕМЕННЫЕ АНАЛИТИЧЕСКИЕ ПРИЛОЖЕНИЯ КЛАССА DATA VISUALIZATION

Современные аналитические приложения класса Data Visualization обеспечивают:

- Быструю скорость обработки больших данных и их ввода-вывода.
- Самостоятельное конфигурирование системы пользователями с использованием метода Drag and Drop и возможностью гибкого определения измерений, иерархий, групп и различных наборов данных.
- Обширный функционал для визуализации данных. Помимо создания различных средств для создания диаграмм, информационных панелей и других пользовательских объектов для аналитики, подразумевается возможность взаимодействия данных и их графическим представлением и наоборот, то есть реализуется высокая степень интерактивности между пользователями и данными.
- Своевременность и актуальность обновления данных и отчетов.
- Возможность интегрировать данные из разных источников, наличие API для соединения с другими сервисами и веб-подключение.

Задачи, решаемые BI-приложением для обеспечения процесса принятия управленческих решений:

- Возможность моделирования бизнес-ситуаций.
- Мониторинг оперативных показателей в каждом бизнес-процессе и их оценка (эффективность использования ресурсов, показатели финансовой и инвестиционной деятельности, показатели службы маркетинга, взаимосвязь операционных показателей со стратегией), а также их анализ по нестандартным запросам.
- Возможность использования глубокой аналитики данных.
- Устранение повторяющихся действий по обработке информации за счет автоматизации и масштабирования.
- Планирование деятельности.
- Оптимизация работы с данными компании.

BI-продукты позволяют использовать инструменты для анализа данных с привычным интерфейсом.

II. OLAP-ТЕХНОЛОГИИ

Типовое хранилище данных, как правило, отличается от обычной реляционной базы данных. Во-первых, обычные базы данных предназначены для того, чтобы помочь пользователям выполнять повседневную работу, тогда как хранилища данных предназначены для принятия решений.

В значительной степени проблемы обработки данных на основе классического подхода решают переход к объединению и анализу данных учетной системы с помощью технологии Online Analytical Processing (OLAP). OLAP-моделью называется логическая модель данных, лежащая в основе OLAP-технологии. Использование технологии OLAP для обработки запросов позволяет повысить скорость и удобство обработки данных. Клиент-серверная архитектура OLAP-продуктов обеспечивает одновременный доступ большого числа пользователей, при этом анализ проводится одинаково быстро по всем аспектам информации независимо от размера и сложности структуры базы данных. Существуют три типа OLAP:

- многомерная OLAP (Multidimensional OLAP – MOLAP);
- реляционная OLAP (Relational OLAP – ROLAP);
- гибридная OLAP (Hybrid OLAP – HOLAP).

Сложность в применении OLAP состоит в создании запросов, выборе базовых данных и разработке схемы, в результате чего большинство современных продуктов OLAP поставляются вместе с предварительно настроенными запросами. Другая проблема – в базовых данных. Они должны быть полными и непротиворечивыми.

С точки зрения пользователя, все варианты выглядят похожими по возможностям. Наибольшее применение OLAP находит в продуктах для финансового планирования, хранилищах данных, решениях класса Business Intelligence.

Полномасштабная информационно-аналитическая система должна выполнять сложные и разнообразные функции, включающие сбор данных из различных источников, их согласование, преобразование и загрузку в хранилище, хранение аналитической информации, регламентную отчетность, поддержку произвольных запросов, многомерный анализ. Информационно-аналитические системы направлены на комплексную информационно-аналитическую поддержку принятия решений.

III. Выводы

Логическая структура данных хранилища данных существенно отличается от структуры данных источников данных. Для разработки эффективного процесса преобразования необходима хорошо проработанная модель корпоративных данных и модель технологии принятия решений. Данные для пользователя удобно представлять в многомерных базах данных.

Хранилища данных содержат интегрированные данные. Они интегрированы на множестве уровней: на уровне ключа, атрибута, на описательном, структурном уровне и так далее. Общие данные и общая обработка данных консолидированы и являются единообразными для всех данных, которые подобны или схожи в хранилище данных.

Кроме извлечения данных из базы, для принятия решений важен процесс извлечения знаний, в соответствии с информационными потребностями пользователя. С точки зрения пользователя в процессе извлечения знаний из базы данных должны решаться следующие преобразования: данные → информация → знания → полученные решения.

Внедрение хранилища данных в работу предприятий позволяет увеличить скорость принятия обоснованных решений с учетом автоматизированной обработки информации, эффективно использовать не только финансовые, но и человеческие и материальные ресурсы компании.

СПИСОК ЛИТЕРАТУРЫ

1. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных OLAP и DataMining, 2004. гл. 4, 5, 7.
2. Ralf Kimball, Margy Ross. The Data Warehouse. 2-Edition. N.Y.: John Wiley, 2002.
3. Ильяшенко, О. Ю. Роль BI-систем в совершенствовании процессов обработки и анализа бизнес информации / И. В. Ильин// Наука и бизнес: пути развития – 2017. –с. 124–131.
4. Андерсон, К. Аналитическая культура. От сбора данных до бизнес-результатов. / М.: Манн, Иванов и Фербер// – 2017. –с. 336.
5. Chen, H. C., Chaing, R. H. L., Storey, V. C. Business intelligence and analytics: From big data to big impact.// MIS Quarterly – 2012. –р. 1165–1188.
6. Trieu, V. H. Getting value from Business Intelligence systems: A review and research agenda. Decision Support Systems – 2017. –р. 111–124.

ЧУВСТВИТЕЛЬНОСТЬ ФИЛЬТРОВ С ЭЛЕМЕНТАМИ ВЫСШЕГО ПОРЯДКА

Петровский И. И., Свито И. Л., Нехаичик Е. В.

Кафедра теоретических основ электротехники,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {Nehajchik, Svito}@bsuir.by, Petrovski47@mail.ru

В работе рассматривается чувствительность фильтрующих устройств, построенных с использованием элементов высшего порядка в зависимости от параметров элементов в структурах фильтров.

ВВЕДЕНИЕ

Применение элементов высшего порядка в электрических фильтрах обеспечивает требуемые амплитудно- и фазочастотные характеристики фильтрующих устройств [1,2].

I. ЧУВСТВИТЕЛЬНОСТЬ ЭЛЕКТРИЧЕСКИХ ФИЛЬТРОВ

Так как элементы высшего порядка построены на активных элементах, в частности, на операционных усилителях, то возникает вопрос устойчивости работы и, соответственно, чувствительности электрических фильтров к изменению параметров элементов [3].

Математическая модель любой электрической цепи может быть представлена в виде передаточной функции:

$$W(p) = \frac{A(p)}{B(p)}$$

При этом, для исследования влияния какого-либо параметра на передаточную функцию, её можно представить в виде:

$$W(p) = \frac{A_1(p) + kA_2(p)}{B_1(p) + kB_2(p)}$$

где k – параметр цепи.

Тогда чувствительность электрической цепи можно рассчитать по формуле:

$$S_k^w = \frac{\delta W}{\delta k}$$

В теории фильтров также исследуется чувствительность нулей и полюсов. При этом можно также рассчитать чувствительность добротности и резонансной частоты в зависимости от параметров элементов.

РАСЧЁТ ЧУВСТВИТЕЛЬНОСТИ

Элементы высшего порядка содержат в своей структуре как активные, так и пассивные элементы, которые влияют на характеристики фильтра. Для дальнейшего исследования возьмём схемы фильтров с элементами второго порядка, представленные на рис. 1а, 1б [1].

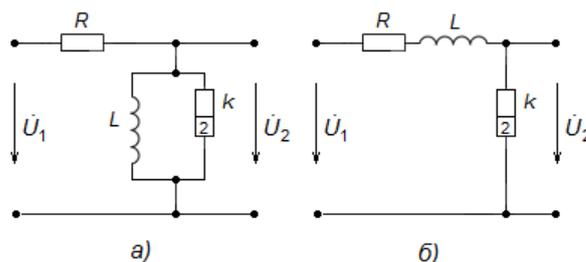


Рис. 1 – Схемы фильтров

Передаточные функции этих фильтров, соответственно будут:

$$W(p) = \frac{p^2}{p^2 + \frac{R}{L}p + \frac{R}{k}}, W(p) = \frac{p^2}{p^2 + \frac{L}{k}p + \frac{R}{k}}$$

С учётом внутренней структуры элементов высшего порядка передаточная функция фильтра (рис. 1,а) имеет вид [4]

$$W(p) = \frac{R_m R_4 T_\alpha^2 p^2}{R_m R_4 T_\alpha^2 p^2 + 4R_3 R \tau p + R_3 R}$$

После преобразования эту функцию можно представить:

$$|W(\omega)| = \sqrt{\frac{R_m^2 R_4^2 T_\alpha^4 \omega^4}{R_m^2 R_4^2 T_\alpha^4 - 2R_m R_4 T_\alpha^2 R R_3 \omega^2 + 16R_3 R^2 \tau^2 \omega^2 + R^2 R_3^2}}$$

Чувствительность функции фильтра от различных параметров цепи после нормирования $x = \frac{\omega}{\omega_0}$

$$S_{R_4}^{|W(\omega)|} = S_{R_m}^{|W(\omega)|} = 1 - \frac{x^4 - x^2}{x^4 - 1,84x^2 + 1},$$

$$S_{\tau_\alpha}^{|W(\omega)|} = 2 - \frac{2x^4 - 2x^2}{x^4 - 1,84x^2 + 1},$$

$$S_{R_3}^{|W(\omega)|} = \frac{0,84x^2 - 1}{x^4 - 1,84x^2 + 1},$$

$$S_\tau^{|W(\omega)|} = \frac{0,16x^2}{x^4 + 1,34x^2 + 1},$$

где значения параметров схемы:

$$R = 100\text{Ohm}, R_3 = 10^3\text{Ohm}, R_4 = 10^3\text{Ohm},$$

$$\tau_\alpha = 10^{-4}\text{c}, \tau = 10^{-5}\text{c}, \omega_0 = 10^4\text{c}^{-1},$$

Графически эти характеристики чувствительности представлены на рис.2.

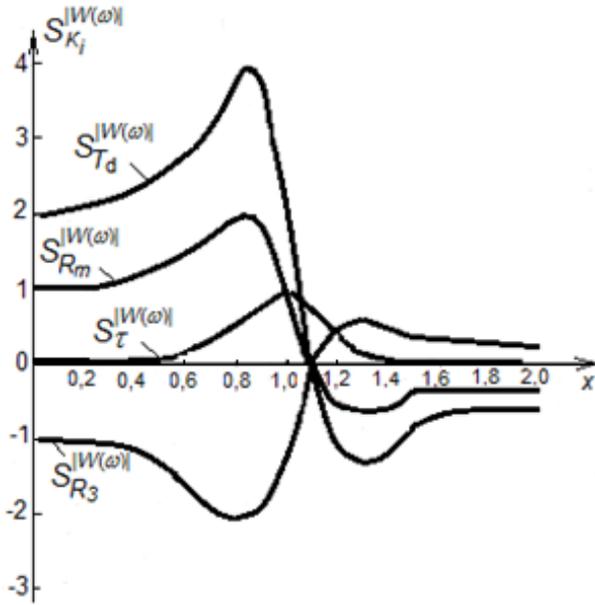


Рис. 2 – Чувствительность фильтра, представленного на рис. 1,а

В соответствии с ранее изложенным, передаточная функция фильтра (рис. 1.б) будет иметь вид:

$$W(p) = \frac{R_L T_i^2 \tau p^2}{R_L T_i^2 \tau p^2 + 4R_L T_i^2 p + R_T}$$

Функции чувствительности от параметров элементов и после нормирования имеют вид:

$$S_{R_L}^{|W(\omega)|} = 1 - \frac{0,64x^4 - 0,26x^2}{0,64x^4 - 0,96x^2 + 1},$$

$$S_{T_i}^{|W(\omega)|} = 2 - \frac{1,3 * 10^{-2}x^4 + 6,4 * 10^5 x^2}{6,4 * 10^{-3}x^4 + 6,4 * 10^5 x^2 + 10^{-2}},$$

$$S_R^{|W(\omega)|} = \frac{0,9x^2 - 1}{0,64x^4 - 0,96x^2 + 1},$$

$$S_T^{|W(\omega)|} = 1 - \frac{0,64x^4 - 1,6x^2 + 1}{0,64x^4 - 0,96x^2 + 1},$$

где значения параметров схемы:

$$R = 500\text{Ohm}, R_3 = 500\text{Ohm},$$

$$\tau_L = 10^{-4}\text{c}, \tau = 10^{-3}\text{c}, \omega_0 = 4 * 10^3\text{c}^{-1},$$

Графически эти характеристики чувствительности представлены на рис. 3.

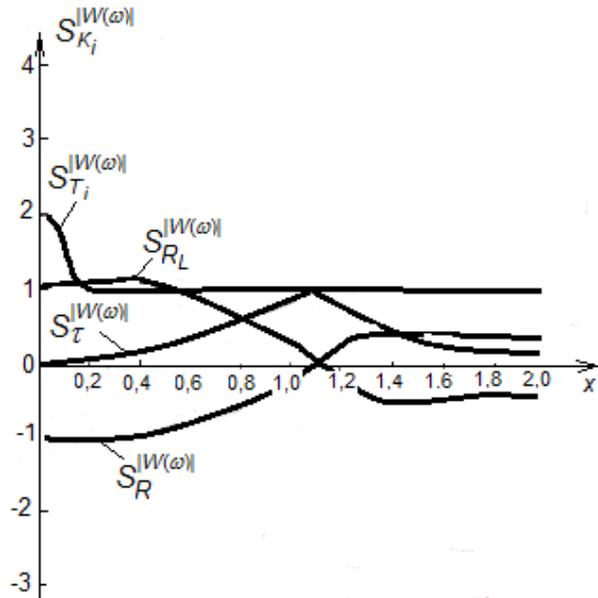


Рис. 3 – Чувствительность фильтра, представленного на рис. 1,б

II. ЗАКЛЮЧЕНИЕ

Сравнительный анализ чувствительности традиционных активных фильтров на гираторах и операционных усилителях с фильтрами на элементах высшего порядка показывает, что чувствительность последних не выше в сравнении с классическими активными фильтрами.

СПИСОК ЛИТЕРАТУРЫ

1. И. И. Петровский, И. Л. Свито. Применение элементов высшего порядка в фильтрах нижних и высших частот. Материалы международной научной конференции «ИТС-2020», 2020 г., БГУИР Минск, Беларусь.
2. И. И. Петровский, И. Л. Свито, Л. Ю. Шилин. Элементы высшего порядка в полосовых фильтрах. Материалы международной научной конференции «ИТС-2021», 2021 г., БГУИР Минск, Беларусь.
3. А. Е. Знаменский, И. Н. Теплоу. Активные RC – фильтры. - М.: Связь, 1970.
4. Petrovski, I. Beitrag zum Entwurf aktiver Spannungsteiler Filter mit Elementen hoherer Ordnung. Dissertation, TH Ilmenau, DDR, 1982.

СПОСОБ УПРЕЖДАЮЩЕГО ТОКООГРАНИЧЕНИЯ ЭЛЕКТРОМАШИНЫ

Сидоров Д., Марков А. В.

Кафедра систем управления

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: sam65th@mail.ru, markov@bsuir.by

В работе рассмотрен и проанализирован способ упреждающего токоограничения электродвигателя (ЭД), ориентированный на применение в позиционных системах автоматического управления (САУ) ЭД.

ВВЕДЕНИЕ

В большинстве позиционных САУ электроприводов (ЭП) используется принцип подчиненного регулирования, появившийся в семидесятих годах прошлого века и не теряющий своей актуальности ввиду простоты своей настройки.

На этапе синтеза контура тока САУ ЭП, зачастую, на модульный оптимум [1], принимают следующие допущения:

1. ротор ЭД не вращается либо заторможен;
2. все малые задержки, вносящиеся аperiodическими звеньями с постоянными времени, соответствующими задержкам датчика тока, фильтра датчика тока, нелинейной по своему характеру временной задержке инвертора, учитываются одним аperiodическим звеном первого порядка с суммарной постоянной времени этих некомпенсированных инерционностей контура регулирования тока [1];

а также используют в качестве регулятора тока (РТ) ПИ-регулятор тока для устранения статической ошибки по моменту нагрузки.

I. ПРОБЛЕМА ИССЛЕДОВАНИЯ

Усложнением структурной схемы контура тока является проявление внутренней ОС по ЭДС двигателя (рисунок 2) при ненулевой скорости вращения двигателя.

При наличии ЭДС вращения ЭД выходной сигнал ПИ-регулятора тока после быстрого завершения регулирования непосредственно тока начинает компенсировать возрастающий сигнал ЭДС (рисунок 1).

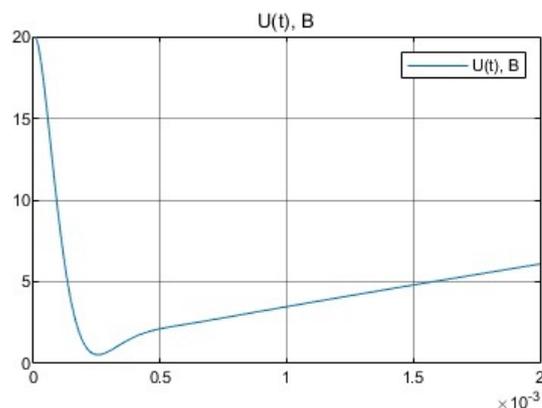


Рис. 1 – Переходный процесс на выходе РТ без компенсации ЭДС вращения ЭД

Таким образом, допущение п.1 является неадекватным в случае вращающегося ротора ЭД. Для устранения данной неадекватности модели контура регулирования тока применяется способ упреждающего токоограничения электродвигателя, учитывающий влияние ЭДС вращения ЭД.

II. УЧЕТ ЭДС ВРАЩЕНИЯ ЭЛЕКТРОМАШИНЫ В КОНТУРЕ ТОКА

Для практической реализации контура управления тока как при заторможенном, так и при вращающемся роторе, требуется учитывать, что значения тока должны быть ограничены в технически допустимых для конкретного ЭД пределах. Это обеспечивается установлением ограничения на выходе регулятора скорости (ограничением задания тока) совместно с установлением ограничения задания напряжения на выходе регулятора тока. Последнее называется «упреждающим токоограничением» [2].

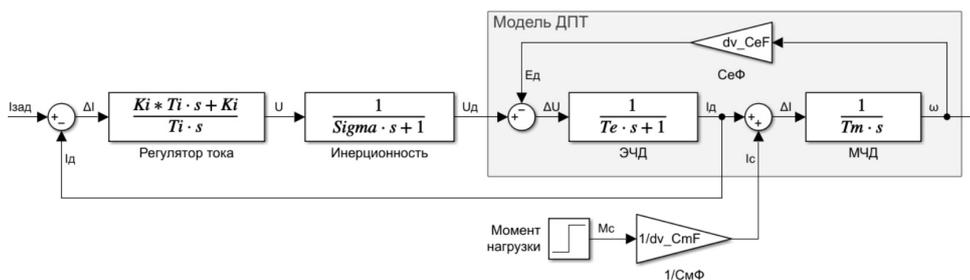


Рис. 2 – Структурная схема контура тока с моделью двигателя постоянного тока независимого возбуждения (ДПТ НВ)

Для реализации упреждающего токоограничения помимо нелинейного звена «насыщения» на выходе регулятора тока (РТ) требуется учесть ЭДС вращения ЭД. Для этого к выходному сигналу РТ добавляется сигнал ЭДС ЭД, который может быть получен с датчика ОС по скорости (тахогенератора, энкодера и т.д.) или непосредственно с датчика скорости, имея в виду, что САУ настроена, и фактическая скорость ЭД слабо отличается от ее задания. Последний вариант является наиболее предпочтительным с точки зрения помехозащищенности всей САУ, но может привести к незначительному уменьшению общего быстродействия системы. Выбор способа учета ЭДС зависит от конкретных особенностей (мощность, уровень помех, требуемая точность, экономическая составляющая и т.д.) позиционно-го ЭП.

Смысл упреждающего токоограничения заключается в том, что управляющий сигнал РТ, соответствующий заданному допустимому уровню тока, подается суммарно с сигналом, равным ЭДС двигателя.

К любому возможному значению ЭДС E (соответствующей ω) добавляется некоторое предельное значение $\pm \Delta U_{max} = IR + L \frac{di}{dt}$ с выхода РТ, которое не может вызвать ток $\pm \Delta I_{max}$ в ЭД, превышающий предельное значение в установившемся режиме:

$$\pm \Delta I_{max} = \frac{\pm \Delta U_{max}}{R} \quad (1)$$

Так, настройка звеньев насыщения регуляторов скорости (РС) и тока приобретает понятный физический смысл: на выходе РС звено насыщения ограничивает максимально допустимое значение задания тока, а на выходе РТ звено насыщения ограничивает максимально допустимое приращение напряжения к ЭДС вращения ЭД, которая пропорциональна сопротивлению статора, автоматически ограничивая напряжение питания двигателя.

На рисунке 3 показано, что узел токоограничения реализуется в виде ПОС по ЭДС двигателя, шунтирующей вход управления инвертора.

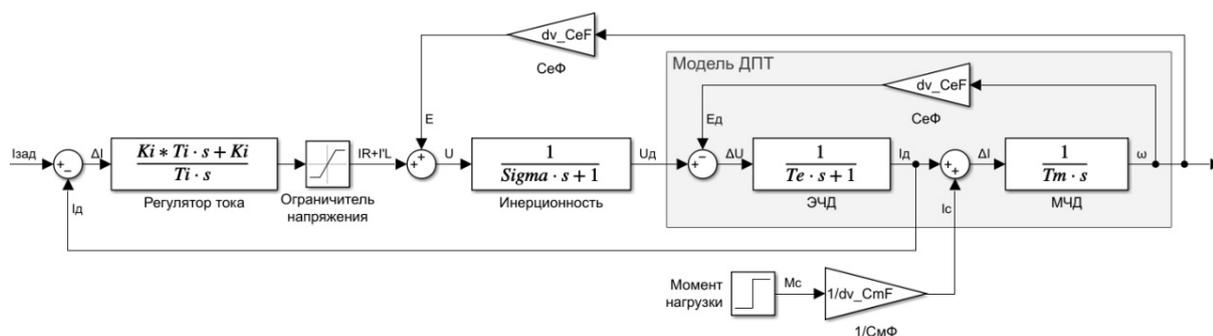


Рис. 3 – Структурная схема контура тока с моделью ДПТ НВ с упреждающим токоограничением

На рисунке 4 показано, что в системе с упреждающим токоограничением переходный процесс с выхода РТ стал иметь сходящийся характер.

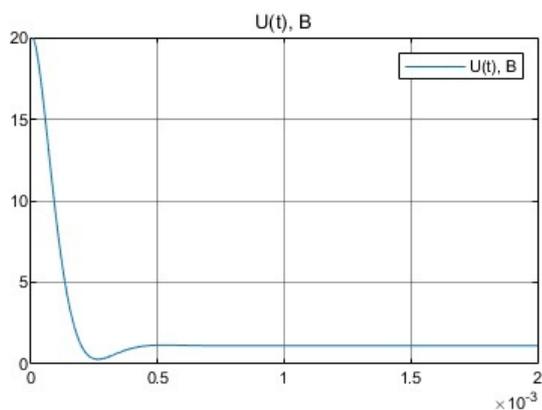


Рис. 4 – Переходной процесс на выходе РТ с компенсацией ЭДС вращения ЭД

III. ЗАКЛЮЧЕНИЕ

Таким образом, был рассмотрен и проанализирован способ упреждающего токоограничения ЭД на отлаженной модели контура регулирования тока САУ ДПТ НВ, настроенной на технический оптимум [1]. Помимо линеаризации контура тока, применение данной ОС обеспечивает меньший диапазон работы РТ, что повышает его точность и быстродействие, что положительно сказывается на точности и быстродействии системы в целом.

1. Фрер, Ф. Введение в электронную технику регулирования / Ф. Фрер, Ф. Отгенбургер, – М. : Энергия, 1974. – С. 127-142.
2. Семко, И. А. Статический расчет систем автоматизированного электропривода постоянного тока – учебное пособие, 2016. – С. 61-62.
3. Анхимок, В. Л., Ильин, О. П. Проектирование систем автоматического управления электроприводами – Мн. : «Вышэйшая школа», 1971. – С. 38-42.

АНАЛИЗ КОНСТРУКЦИИ И ПАРАМЕТРОВ СИСТЕМЫ МЕЖСОЕДИНЕНИЙ ПЕЧАТНОЙ ПЛАТЫ

Стещенко П. П., Журавлев В. И.
Кафедра теоретические основы электротехники,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: sppmailbox@mail.ru, vadzh@bsuir.by

Проведен анализ линий передачи сигналов, предложены конструкция и расчет электрофизических параметров системы межсоединений печатной платы.

ВВЕДЕНИЕ

Печатные платы являются составной частью современной радиоэлектронной аппаратуры (РЭА). Они выполняют важную роль в интегрированном исполнении конструкций РЭА. Полупроводниковые устройства, смонтированные на печатной плате, представляют сложные функциональные устройства для обработки большого массива информации.

1. ТРЕБОВАНИЯ К КОНСТРУКЦИИ

Для обеспечения быстродействия полупроводниковые элементы формируются с малыми субмикронными размерами ($\sim 10^{-8}$ м) и с малыми уровнями сигналов ($\sim 10^{-3}$ А,В). Печатные платы, на которых устанавливаются полупроводниковые устройства служат для обработки сигналов более высокого уровня ($10^{-3} \div 10^3$ А,В) необходимые для управления исполнительными элементами (электромагнитные реле, полупроводниковые ключи, электродвигатели и т.д.).

Указанные параметры предопределяют увеличение геометрических размеров элементов межсоединений печатной платы (длина, толщина линий соединений и расстояния между ними), необходимых для обеспечения допустимой плотности тока в них и диэлектрической изоляции между ними. В результате увеличиваются емкости межсоединений (снижение быстродействия) и тепловая нагрузка на печатную плату (необходимость охлаждения).

Поэтому при разработке конструкции печатной платы с заданными параметрами необходимо решить ряд важных проблем:

- определить наиболее важные функциональные элементы и их размещение на печатной плате;
- оптимизировать размеры элементов межсоединений для получения заданных электрофизических параметров;
- рассчитать их тепловое воздействие на окружающие элементы;
- моделировать оптимальную компоновку (топологию) для интегрированной конструкции;
- уменьшить влияние окружающей среды на надежность конструкции.

При решении указанных проблем необходимо применять (или разрабатывать) математические и физические модели элементов интегрированных в единой конструкции радиоэлектронных элементов. Такие модели должны учитывать схемотехнические решения устройства, свойства применяемых материалов, климатические зоны их эксплуатации, прогнозировать заданные параметры надежности. Сложность комплексного решения указанных проблем связана с различием физических, математических и технических моделей элементов печатной платы. Поэтому в настоящее время разработка печатной платы проводится методом отдельной оптимизации параметров по каждому виду модели, что затрудняет создание единой конструкции РЭА с заданными параметрами.

Нами рассмотрена одна из проблем: оптимизация размеров элементов межсоединений для обеспечения заданных их электрофизических параметров.

Межэлементные соединения печатной платы представляет собой линии передачи сигналов, которые рассчитываются с учетом степени интеграции схемы и частотных параметров. В качестве соединений используются в основном конструкции с симметричными, несимметричными и компланарными линиями [1].

В несимметричной линии передачи часть электромагнитной энергии находится в пространстве над проводником. Этот воздушный зазор определяет ограничения по частоте и мощности, так как возникают трудности согласования сопротивлений открытого пространства и линии передачи. Наиболее близкой к конструкции печатной платы, по нашему мнению, подходит симметричная линия передачи с «подвешенной» подложкой показанной на рис. 1.

Важными параметрами линий передачи являются:

- рабочая частота,
- волновое сопротивление
- и затухание.

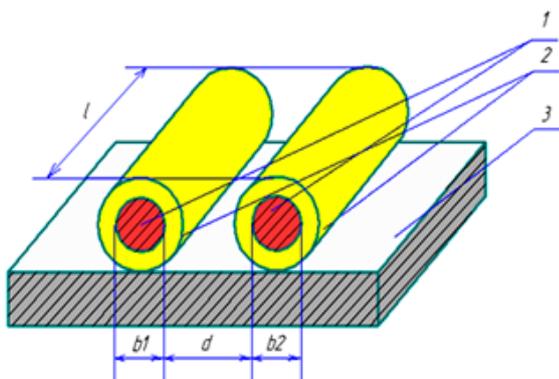


Рис. 1 – Симметричная линия передачи с «подвешенной» подложкой

Рабочая частота f_c должна быть ниже критической частоты, определяемой по формуле:

$$f_c = \frac{75}{h \cdot (\varepsilon - 1)^{\frac{1}{2}}},$$

где h – толщина изоляции (поз.2, рис. 1), ε – диэлектрическая проницаемость изоляции.

Волновое сопротивление линии Z_b определяется из соотношения:

$$Z_b = \frac{120}{\sqrt{\varepsilon}} \ln \frac{a - r}{r},$$

где a – половина расстояния между проводниками.

С увеличением диэлектрической проницаемости [25В?] и радиуса r (площади поперечного сечения) проводника волновое сопротивление уменьшается.

В линиях передачи сигнала затухание складывается из трех частей: затухание, определяемое

потерями в проводнике, обусловленное потерями в диэлектрике, затухание вследствие излучения. В линиях передачи затуханием вследствие излучения ввиду его малости можно пренебречь. При использовании подложки с высокой диэлектрической проницаемостью (керамика) потери в диэлектрике можно также не учитывать.

При выборе материала проводников в первую очередь принимают во внимание их сопротивление на высокой частоте, коэффициент термического расширения. Наиболее широко используются следующие материалы: медь, серебро, золото, алюминий, хром, тантал, титан. Электрофизические свойства материалов представлены в справочной литературе [2]. Выбор материала межсоединений проводится согласно требований к конструкции печатной платы и условий ее эксплуатации.

Приведенный анализ конструкции, параметров системы межсоединений печатной платы может использоваться студентами по специальности «Промышленная электроника» при выполнении дипломного проектирования для разработки конструкции печатной платы РЭА [3].

СПИСОК ЛИТЕРАТУРЫ

1. Гроднев И.И. Линии связи. Учебник для вузов. /И. И. Гроднев. С..М. Верник.//Москва. Радио и связь.1988.
2. Кошкин Н.И. Справочник по элементарной физике./ Н.И. Кошкин,М.Г. Ширкевич.//Москва. Государственное издательство физико-математической литературы.1962.
3. Стешенко П.П. Применение адаптивных методов в изучении технических дисциплин. П.П. Стешенко, В.И. Журавлев. //Международная научно-практическая конференция «Качество образовательного процесса: проблемы и пути развития» Минск. БГУИР 26 апреля 2022г с.71-72

СРАВНЕНИЕ ЭФФЕКТИВНОСТИ РЕГУЛЯТОРОВ ЭЛЕКТРОПРИВОДА ГРУЗОВОЙ ТЕЛЕЖКИ ПРИ ПОДАВЛЕНИИ КОЛЕБАНИЙ ПОДВЕСА

Хаджинов М. К., Павлова А. В., Стасевич Н. А.

Кафедра систем управления, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {pavlova, stasevich}@bsuir.by

Решается задача управления краном с целью подавления колебаний подвеса с грузом при перемещении электропривода грузовой тележки в горизонтальном направлении. Производится сравнение эффективности алгоритмов управления электроприводом при подавлении колебаний подвеса.

ВВЕДЕНИЕ

В простейшем случае движение груза и тележки описывается линейной двухмассовой моделью третьего порядка с переменными пространства состояний: скоростью тележки, скоростью груза и отклонением координаты груза от точки подвеса грузовой тележки [1]. Управление осуществляется через электропривод грузовой тележки, имеющий высокие динамические характеристики и полосу пропускания, во много раз превышающую частоту собственных колебаний подвеса груза. Ветровые возмущения приложены к грузу в виде дополнительного ускорения.

1. СРАВНЕНИЕ ЭФФЕКТИВНОСТИ ЭЛЕКТРОПРИВОДОВ

Демпфирования колебаний подвеса и действия возмущений возможны в замкнутом контуре управления с дополнительными датчиками: скорости груза или отклонения груза. Сигналы эквивалентные дополнительным датчикам могут быть сформированы наблюдателем, привязанным к установленному на электроприводе датчику скорости тележки.

Сравним эффективность привода грузовой тележки на примере реализации следующих алгоритмов:

- с перенастройкой регулятора контура скорости;
- без перенастройки регулятора контура скорости с использованием стандартного ПИД-регулятора;
- с модальным регулятором контура оценивания модели крана.

В первом случае канал передачи возмущений от раскачивания груза на выход электропривода грузовой тележки может регулироваться изменением частоты среза контура скорости электропривода. Уменьшением частоты среза контура скорости электропривода и его способности удерживать грузovou тележку от колебаний можно достичь желаемого демпфирования колебаний подвеса с грузом [2].

Для сравнения эффективности рассмотрим алгоритм демпфирования колебаний подвеса крана от датчика скорости электропривода грузовой

тележки, как по управлению, так и по возмущению без перенастройки частота среза электропривода, т.е. со стандартным приводом грузовой тележки, к которому относится ПИД-регулятор.

Нули ПИД регулятора настраиваются на колебательные полюсы подвеса, подъём дифференциальной составляющей ПИД регулятора ограничиваем квадратом частоты колебаний подвеса. На рис. 1 представлены частотные характеристики ПИД-регулятора.

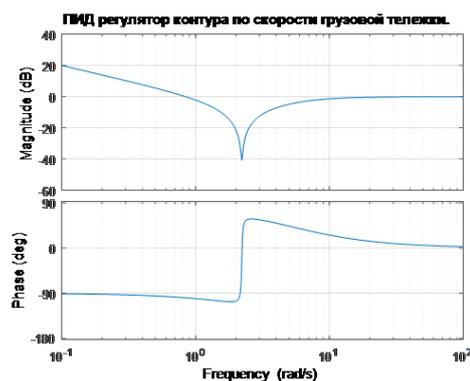


Рис. 1 – Частотные характеристики ПИД-регулятора

Для демпфирования колебаний по возмущению стандартным приводом грузовой тележки в регуляторе должна присутствовать модель крана, которая может привязываться к датчику скорости грузовой тележки. Частота крана и динамика подвеса остается неизменной, увеличим демпфирование подвеса парой полюсов.

При моделировании заданы следующие параметры крана, его нагрузка и настройки: длина подвеса массы на блоке – $L = 2$ м, частота среза электропривода грузовой тележки – 1000 1/с, желаемый коэффициент демпфирования подвеса – 0,75.

При моделировании учитывалось влияние соотношения между массой груза и массой тележки на динамику процессов по отклонению и возмущению.

При отношении массы груза к массе грузовой тележки равное 10, т.е. когда кран перегружен процессы представлены на рис. 2.

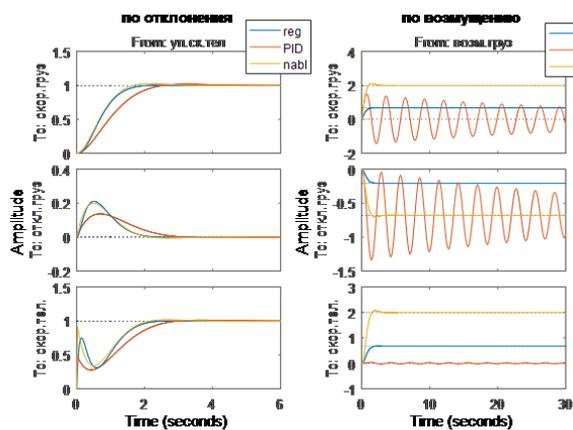


Рис. 2 – Графики процессов

При отношении массы груза к массе грузовой тележки равное 1, т.е. когда кран нормально нагружен, процессы представлены на рис. 3.

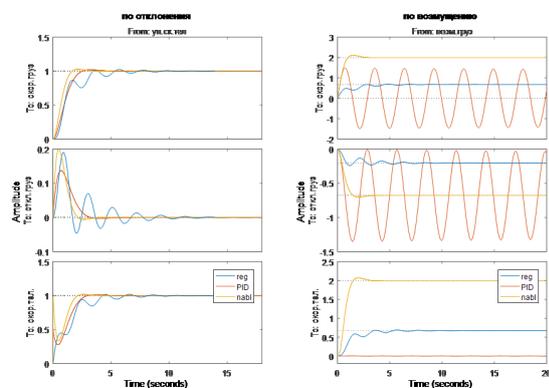


Рис. 3 – Графики процессов

При отношении массы груза к массе грузовой тележки равное 0,1, т.е. когда кран не нагружен, процессы представлены на рис. 4.

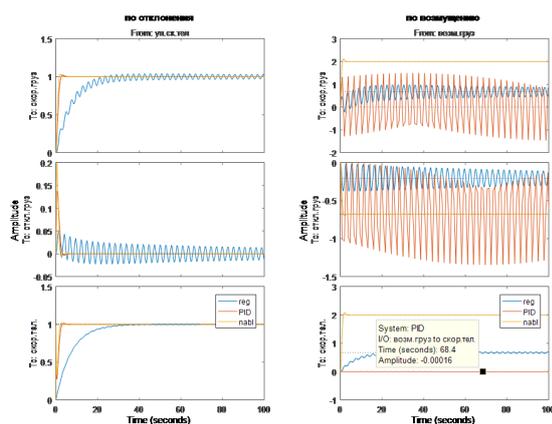


Рис. 4 – Графики процессов

При нормальной нагрузке крана перенастройка регулятора контура скорости обеспечива-

ет превосходное демпфирование колебаний подвеса груза как по управлению, так и по возмущению от ветровых нагрузок. При заниженной нагрузке крана перенастройка регулятора контура скорости обеспечивает в какой то степени приемлемое демпфирование колебаний подвеса груза как по управлению, так и по возмущению от ветровых нагрузок.

При нормальной и заниженной нагрузке крана регулятор обеспечивает превосходное демпфирование колебаний по управлению. Длительность процессов не превышает 3 с. По возмущению ситуация противоположная. Контур регулирования от датчика скорости электропривода ужесточают характеристики электропривода и ещё больше уменьшают его малые демпфирующие свойства.

При нормальной и заниженной нагрузке крана наблюдатель обеспечивает превосходное демпфирование колебаний по управлению и по возмущению. Также эффективно демпфируются колебания и не нагруженного подвеса.

II. ЗАКЛЮЧЕНИЕ

Под термином «демпфирование колебаний подвеса груза» обычно понимают два разных физических явления сходных по практическому результату:

- процесс управления с не возбуждением свободных колебаний подвеса груза
- процесс подавления колебаний подвеса груза.

ПИД-регулятор с настройкой нулей на колебательные полюсы подвеса хотя успешно решает первую задачу, но одновременно, за счёт интегральной составляющей, всячески препятствует решению второй.

Таким образом, любые интегральные регуляторы в контуре от датчика скорости грузовой тележки или позиционный электропривод не приведут к системе подавления колебаний подвеса груза.

Перспективным представляется использование наблюдателя сигнала отклонения груза, позволяющего эффективно демпфировать колебания и не нагруженного подвеса.

1. Хаджинов, М. К. Система управления подъёмным краном на базе квазимодального регулятора с функцией подавления колебаний перемещаемого груза / М. К. Хаджинов, А. С. Шмарловский // Доклады БГУИР. – 2009. – № 7. – С. 38–43.
2. Хаджинов, М. К. Подавление колебаний подвеса крана настройкой регуляторов электропривода грузовой тележки / М. К. Хаджинов, А. В. Павлова, Н. А. Стасевич // Материалы международной научной конференции «Информационные технологии и системы» (ИТС 2022). – 2022.

INTELLIGENT STORAGE AND ACCOUNTING SYSTEM

Varchenya I. V., Savenko A. G.

Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
E-mail: savenko@bsuir.by

The paper presents a software and hardware complex for automation and intellectualization of storage and inventory control of goods at an enterprise, implemented on the Raspberry Pi hardware platform with voice control and visual light indication for request processing. The developed software and hardware system for storage and inventory control can be easily scaled and adapted to the specifics and needs of various enterprises and will automate and facilitate this process.

INTRODUCTION

In the modern world, there are many solutions that automate work in production, agriculture and everyday life. All such solutions are controlled using a specific list of commands hidden from the average user under the interfaces of various mobile and computer applications.

This paper presents a developed project for the automation and intellectualization of a system for storing and inventorying goods, which includes developed software and hardware and a prototype of the system. The prototype of the proposed system is intended for storage and inventory control of radio components at an enterprise engaged in the repair of electrical equipment. The system itself is easily scalable by appropriately changing the system hardware (sensors, sensors appropriate for the task, etc.).

I. PROPOSED SOLUTION

The smart storage and inventory control system is implemented in hardware on the basis of a Raspberry Pi microcomputer, which allows you to economically and ergonomically build the structure of the physical part of the project to manage the system. The Raspberry Pi is a single-board microcomputer, originally developed as a low-cost system for teaching computer science, which subsequently became much more widely used and popular than its developers expected [1].

Peripheral devices are connected to the Raspberry Pi microcomputer, such as: weight sensors (to determine the presence and quantity of stored parts), LED display devices (to visualize found storage containers or display search errors), microphone (for receiving incoming voice commands) and other devices, depending on the specifics of the storage system. The strain gauge and the HX711 chip are a combination that will allow you to create devices for measuring weight or pressure exerted on the surface of the sensor, and then transmit these readings (with high accuracy) to a microcomputer board. For light indication of search results (a specific rack or container) based on a voice search query (in green), as well as for indicating the absence of search results (in red), it is proposed to use LED display devices.

An example of a diagram for connecting LED display devices for a rack with sixteen storage containers placed on it (the developed physical model of the prototype) is shown in Fig. 1.

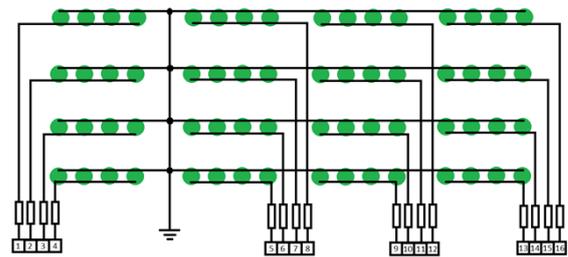


Figure 1 – Example of connecting LED display of a rack

For software control of the system, a software tool has been developed that implements control via a voice interface and does not require special training of employees. The developed software tool uses a number of technologies that implement the following functions:

- processing voice commands;
- processing information read from the weight sensor;
- processing information about placed parts;
- data output via LED indication.

This functionality solves the following tasks:

- voice control software;
- storing information about placed parts;
- input of weight sensor data;
- outputting data to indicator LEDs;
- taking into account operating modes with the physical model.

The general operating algorithm of the system management software involves receiving commands to search for stored parts, add parts for storage, and remove (remove) parts from storage. After starting the program, the LED on the physical model of the system lights up, which serves as an indicator of the operation of the software. Next, the user needs to say the control command out loud.

Control commands are implemented as follows: after the command is read by the microphone, the voice is converted into text. Voice commands are sent to the voice command input unit. At this stage, voice commands represent vibrations of the air environment, perceived by the microphone, pro-

duced by the user. Next, the signal enters the unit for analyzing the received signals from the microphone. Afterwards, the text string is divided into components in the voice command processing unit, according to the list of possible voice commands, and a command is issued to change the entry in the storage file. If the command is recognized, then the corresponding processing method is launched: searching, deleting, adding and changing the number of parts. If the command is not recognized, an error indication is displayed. Information from the weight sensor is analyzed, and then the activation indication command is set in the corresponding block. The indication command generates an error indication command or an indication of a specified cell in the corresponding blocks. Next, activation, error, and specified cell indication commands are sent to the command transmission unit to the microcontroller, which generates control signals on the microcontroller I/O ports. The program ends with a user interrupt. In the case of searching for a part, there can be two outcomes: the part is found and the part is not found. In the first case, the required cell will be highlighted using indicator LEDs. In the second, an error indication will be displayed. In the case of a command to add a new part – the necessary one (with already similar parts or, if there are none, the empty cell will be highlighted. The command to add a part can be of two types: with and without specifying the quantity (then the quantity is considered equal to one). The part removal command can also be with or without specifying the number of radio parts. For example: “Remove voltmeter.” Or: “Remove ten relays.” When you enter a command to change the number of parts, the presence of similar parts is checked. If the check is passed, then the cell with similar details will be highlighted; if not, then the empty cell will be highlighted. The removal command must indicate the number of parts to be removed. In the case of a command to remove a radio component, the corresponding indication of the removal of a part from a certain cell will be displayed and their quantity will be changed. During the development of the software, the following modules were developed, the hierarchical model of which is presented in Fig. 2. The operation of the application begins after the software control module is launched, which contains the code that implements the main algorithm through calls to methods from other modules. The constants module stores unchangeable parameters with information about the name of the information storage file, the values of the pin numbers that control the LEDs, the pin numbers for processing information from the weight sensor, and string values for working with voice commands. The command reader module represents a set of methods that is responsible for reading voice commands directly through the microphone and converting voice commands into

text strings. The text parsing module provides a set of methods that converts text strings into a set of parameters that the program can understand. It also prepares information for recording in an Excel file for storage.

The electronics control module includes two main blocks of code: an LED control unit and a weight sensor reading unit. The LED control unit contains methods for displaying a specific cell, as well as for displaying an indication of an error and removal of a radio component from the cell. The weight sensor reading block contains methods for tracking changes in the weight of the system, as well as a method for processing user clicks to activate the system. The module for working with Excel files contains methods for working with stored information. These are methods for searching, adding, deleting and changing the number of parts.

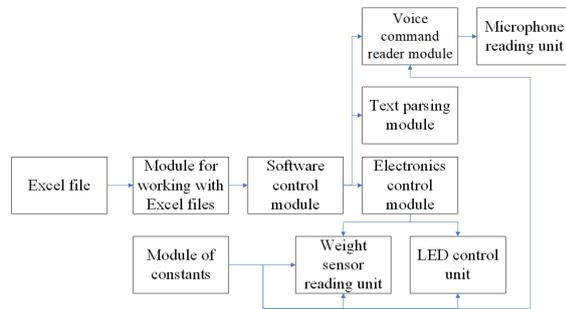


Figure 2 – Software structure

For the correct operation of the application, as well as to fill it with information, a storage file is also used, which is an Excel file consisting of one table with dimensions N by M cells, which corresponds to the dimensions of the physical model of the storage rack, where N is the number of shelves of the rack, M is the number cells (containers) on the shelf of the rack.

Each table cell in an Excel file can contain an unlimited number of records. Each entry is a structure consisting of the name of the stored parts, a colon and a space sign, and a number indicating the number of radio parts. If there are several records in a table cell, they are separated by a line break. All this allows you to use the file not only through voice control of the software, but also in manual mode.

II. CONCLUSION

A software and hardware intelligent system for storage and inventory control of radio components for enterprises has been developed, which can easily be scaled and adapted to the specifics and needs of various enterprises and will automate and facilitate this process.

BIBLIOGRAPHY

1. 3d-diy.ru. [Electronic resource]. – Access mode: <https://3d-diy.ru/wiki/arduino-platy/obzor-plat-raspberry-pi/>. Access date 10.10.2023.

Секция «Интеллектуальные
информационные технологии»

Section «Intelligent information
technologies»

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ МАШИННОГО ОБУЧЕНИЯ В МОБИЛЬНЫХ ПРИЛОЖЕНИЯХ: ОТ ПРОЕКТИРОВАНИЯ ДО РЕАЛИЗАЦИИ

Ахалкина А. М., Федоренко В. А.

Кафедра связи, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: {anastasia.akhapkina2018}@gmail.com

Цель данной работы – изучение областей применения алгоритмов машинного обучения в мобильных приложениях и исследование этапов проектирования мобильного приложения.

ВВЕДЕНИЕ

Недавние преобразования в сфере технологий кардинально изменили нашу жизнь. Ранее мобильные телефоны функционировали исключительно как средства связи. Однако, в недавней истории человечества, эволюция этой технологии достигла таких высот, что телефоны теперь выполняют не только функции коммуникации, но и предоставляют широкий спектр важных и полезных функций.

На данный момент каждый обладатель смартфона использует его не только для общения, но и для управления фитнес-программами, получения образования, выполнения измерений, навигации и многих других значимых задач. Применение алгоритмов искусственного интеллекта, и в частности, технологии машинного обучения (ML), позволяет внедрять в мобильные приложения новейшие инновационные функциональные возможности. В данном контексте, наша статья предлагает анализ областей применения машинного обучения в мобильных приложениях.

I. ОСНОВНАЯ ЧАСТЬ

Сегодня множество компаний стремятся к индивидуализированной коммуникации с клиентами. Эффективность этого подхода демонстрируется с помощью мобильных приложений, специально ориентированных на пользователей. Эти приложения, оснащенные рекомендательными системами, способны собирать и анализировать информацию о пользователях, такую как история покупок и предпочтения товаров в интернет-магазине, а также пользовательские отзывы и оценки. На основе этих данных они могут создавать индивидуализированные рекомендации, предлагая разнообразные товары или предоставляя соответствующий контент [1].

Следующим важным сегментом мобильных приложений можно отметить приложения, способные осуществлять анализ и распознавание изображений, текста и видеоматериалов. Этот раздел искусственного интеллекта широко известен как область **компьютерного зрения**. Технология компьютерного зрения позволяет автоматически распознавать и идентифицировать объекты на фотографиях или в видеозаписях, классифици-

ровать их и, основываясь на обработанных данных, принимать решения без прямого участия человека. Кроме того, разнообразные алгоритмы машинного обучения способны **распознавать человеческую речь** и преобразовывать ее в цифровой формат. Эта функциональность находит применение во многих приложениях, где более удобно взаимодействовать с устройством, используя голосовые команды, а не текстовый ввод. Современные мобильные устройства оборудованы разнообразными сенсорами и датчиками, которые позволяют приложениям собирать информацию об окружающей среде. Алгоритмы машинного обучения способствуют более точной обработке этих данных, повышая качество информации, которая может быть использована в будущем.

Важной технологией также является **автоматизированное рассуждение**, которое дает компьютерам способность применять логические рассуждения для решения разнообразных задач [1]. В результате, машины эмулируют процессы человеческого обучения, анализируя большие объемы данных, извлекая уроки из предыдущего опыта и прогнозируя результаты. Внедрение автоматизированного рассуждения в мобильные приложения позволяет им действовать автономно в решении определенных задач и координации различных процессов.

Однако внедрение всех этих технологий в мобильные приложения представляет собой непростую задачу для разработчиков, и они часто нуждаются в надежной поддержке, предоставляемой библиотеками машинного обучения, встроенными в мобильные операционные системы (ОС). На сегодняшний день существует ряд фреймворков машинного обучения, которые открывают перед разработчиками широкие возможности для создания интеллектуальных мобильных приложений. Эти фреймворки способны обрабатывать обширные объемы данных и предоставляют инструменты для разработки приложений с обработкой данных непосредственно на мобильных устройствах.

Среди библиотек моделей машинного обучения, предназначенных для мобильных приложений, можно выделить следующие: ML Kit, Core ML, TensorFlow Lite и KotlinDL. Каждая из этих библиотек предоставляет разработчикам разно-

образные средства, упрощающие интеграцию технологий машинного обучения в мобильные приложения. Эти библиотеки позволяют выполнять такие задачи, как распознавание текста, лиц, обнаружение и отслеживание объектов, создание меток для изображений, а также разработку собственных моделей классификации. Они также предоставляют функциональность для сканирования штрих-кодов и QR-меток, а некоторые из них обеспечивают возможности идентификации языка, перевода текста на различные языки, а также создания интеллектуальных ответов в текстовых чатах и многих других возможностей. Важным преимуществом данных библиотек является возможность выполнения обработки данных непосредственно на мобильных устройствах, что позволяет использовать их в режиме реального времени.

Кроме того, существуют универсальные библиотеки машинного обучения, которые могут быть успешно применены и в мобильных приложениях. Среди них можно выделить PyTorch, QNNPACK, Keras и ONNX. Эти фреймворки оптимизированы для использования на мобильных устройствах и предоставляют разработчикам мощные инструменты для создания и развертывания моделей машинного обучения. Например, на сегодняшний день PyTorch поддерживает развертывание предварительно обученных моделей на платформах iOS, Android и Linux.

Библиотека QNNPACK, хотя и не предназначена для прямого использования разработчиками приложений с использованием методов машинного обучения, предоставляет низкоуровневые примитивы производительности для высокоуровневых фреймворков глубокого обучения. Напри-

мер, она интегрирована в PyTorch, что позволяет оптимизировать производительность и улучшить качество моделей машинного обучения [2].

II. ЗАКЛЮЧЕНИЕ

Результаты данной работы подчеркивают значительный потенциал и разнообразие применения машинного обучения в мобильных приложениях. Большое количество библиотек моделей машинного обучения, предназначенных для мобильных устройств, открывают перед разработчиками возможности для создания интеллектуальных приложений, способных обрабатывать данные на устройстве пользователя. Процесс проектирования мобильного приложения с использованием алгоритмов машинного обучения охватывает не только технические аспекты, но и функциональные требования и динамические аспекты взаимодействия пользователя с системой.

В целом, данная работа предоставляет обширное представление о возможностях и перспективах применения машинного обучения в мобильных приложениях, а также подчеркивает важность проектирования приложения с учетом алгоритмов машинного обучения для достижения успешных результатов в этой области.

СПИСОК ЛИТЕРАТУРЫ

1. 5 Challenges for Developing Mobile Apps with AI & Machine Learning Capabilities [Electronic resource]. – Mode of access: <https://heartbeat.comet.ml/5-challenges-for-developing-mobile-apps-with-aimachine-learning-capabilities483668704a60>. – Date of access: 12.02.2023. – Title from the screen.
2. QNNPACK [Electronic resource]. – Mode of access: <https://github.com/pytorch/QNNPACK>. – Date of access: 12.02.2023. – Title from the screen.

СТРУКТУРА ИНТЕЛЛЕКТУАЛЬНОЙ ДИАЛОГОВОЙ СИСТЕМЫ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ

Гердий Е. В., Зорин Д. О., Жмырко А. В., Банцевич К. А., Орлов М. К.
Национальный детский технопарк

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

E-mail: {egor.gerdiy, zorindanik, Aleksashazh, ksusha.bantsevich, orlovmaksimkonstit}@gmail.com

В данной статье рассматривается подход к проектированию интеллектуальных диалоговых систем, обеспечивающих эффективные и комфортные условия жизнедеятельности лиц с ограниченными возможностями. Результатом данной работы является структура интеллектуальной диалоговой системы для таких лиц.

ВВЕДЕНИЕ

Интеллектуальные диалоговые системы позволяют обеспечить более эффективное и удобное взаимодействие пользователей с компьютерными системами [1]. В настоящее время задача в разработке таких систем становится всё более актуальной для лиц с ограниченными возможностями. В работе предлагается подход к проектированию интеллектуальной диалоговой системы для постоянной информационной поддержки жизнедеятельности таких лиц [2]. Система должна позволять:

- получать информацию о доступности объектов социальной и транспортной инфраструктуры;
- формировать информацию по наилучшим маршрутам, учитывая потребности таких лиц;
- поддерживать виртуальное общение для снижения уровня социальной изоляции;
- и др.

Работа связана с приоритетными направлениями научно-технической и инновационной деятельности в соответствии со стратегией социально-экономического развития РБ, предусмотренной Государственной программой «Социальная защита» на 2021–2025 годы [3, 4].

I. СТРУКТУРА ПРОЕКТИРУЕМОЙ СИСТЕМЫ

Традиционная структура интеллектуальной системы включает базу знаний, решатель задач и пользовательский интерфейс [5]. В основе интеллектуальных диалоговых систем лежит персонализированный подход к пользователям таких систем. Особенностью рассматриваемых систем является необходимость в обеспечении такого подхода для лиц с ограниченными возможностями.

В качестве инструмента для проектирования таких систем была выбрана Технология OSTIS [6], в рамках которой реализована интеллектуальная диалоговая система НИКА [7]. Данную систему можно легко расширить до интеллекту-

альной диалоговой системы в контексте заданной предметной области.

С точки зрения Технологии OSTIS база знаний представляет собой иерархию предметных областей и соответствующих им онтологий, решатель задач – коллектив взаимодействующих через базу знаний агентов, а пользовательский интерфейс – встраиваемую подсистему для взаимодействия с пользователями [8].

Таким образом, предлагается расширить базу знаний, решатель задач и пользовательский интерфейс системы НИКА для создания интеллектуальной диалоговой системы для лиц с ограниченными возможностями.

II. БАЗА ЗНАНИЙ ПРОЕКТИРУЕМОЙ СИСТЕМЫ

База знаний проектируемой системы должна включать в себя следующие предметные области и онтологии:

- Предметную область и онтологию географических объектов;
- Предметную область и онтологию объектов социальной и транспортной инфраструктуры;
- Предметную область и онтологию пользователей.

Первая предметная область и онтология включает классификацию административно-территориальных объектов, вторая – классификацию социальных объектов и классификацию транспортных сообщений в рамках административно-территориальных объектов, третья – классификацию пользователей, их характеристик (индивидуальных, физиологических, социальных и др.) и действий.

Обработку и наполнение базы знаний выполняет решатель задач проектируемой системы.

III. РЕШАТЕЛЬ ЗАДАЧ ПРОЕКТИРУЕМОЙ СИСТЕМЫ

Решатель задач проектируемой системы должен включать следующие агенты:

- Агент получения информации о географических объектах заданной категории;
 - Агент получения прогноза погоды в заданном географическом объекте;
 - Агент определения местоположения пользователя в рамках заданного географического объекта;
 - Агент построения наилучшего маршрута между заданными географическими объектами;
 - и др.;
- Агент получения информации об объектах социальной и транспортной инфраструктуры заданной категории в рамках заданного географического объекта;
 - Агент получения информации о доступности заданного социального объекта;
 - Агент построения наилучшего маршрута между заданными объектами социальной инфраструктуры;
 - и др.;
- Агент наполнения базы знаний из внешних информационных источников;
 - Интерактивный Telegram-бот для добавления информации о географических объектах, объектах социальной и транспортной инфраструктуры;
- Агент определения предпочтений и возможностей заданного пользователя.

Указанный решатель задач проектируемой системы позволит получать ответы на следующие категории вопросов:

- вопросы о местонахождении заданных объектов социальной и транспортной инфраструктуры (например, «Где находится ближайшая поликлиника?», «Как можно добраться до музея Великой Отечественной войны в городе Минске?», «Где можно получить консультацию по вопросам жилищно-коммунального хозяйства в городе Брест?» и т.д.);
- вопросы о доступности заданных объектов социальной и транспортной инфраструктуры (например, «Оборудована ли ближайшая поликлиника пандусом?», «Есть ли парковка для лиц с ограниченными возможностями в передвижении около ближайшего магазина?» и т.д.);
- вопросы о предпочтениях пользователей относительно заданных объектов социальной и транспортной инфраструктуры (например, «Где можно купить дешёвый обед в городе Минск?», «Какие заведения питания открыты до 21:00 в городе Бобруйск?», «Какие из больниц города Витебск работают круглосуточно?» и т.д.).

Telegram-бот, помимо возможности пополнения информации об объектах социальной и транспортной инфраструктуры, также позволит отме-

чать указать информацию о доступности таких объектов для лиц с ограниченными возможностями.

IV. ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС ПРОЕКТИРУЕМОЙ СИСТЕМЫ

Пользовательский интерфейс проектируемой системы должен включать компонент диалогового окна и компонент интерактивной карты. Диалоговое окно позволит задавать системе вопросы на естественном языке. С помощью карты можно будет не только визуализировать объекты, которые использовались в диалоге с пользователем, но и задать вопросы о выбранных на карте объектах.

В общем случае пользовательский интерфейс системы должен обеспечивать мультимодальность: возможность голосового и текстового ввода и вывода, а также адаптацию под конкретные особенности заданного пользователя [8].

V. ЗАКЛЮЧЕНИЕ

В данной работе была предложена структура интеллектуальной диалоговой системы для лиц с ограниченными возможностями на основе системы НИКА. В дальнейшем планируется реализация указанной системы.

1. Карпов, В. Э. От коллаборативной робототехники к социальным роботам для поддержки людей с ограниченными возможностями: новые направления разработки использования интеллектуальных агентов / В. Э. Карпов, В. Б. Тарасов // Интеллектуальные технологии и средства реабилитации и абилитации людей с ограниченными возможностями (ИТСР-2018). – 2018. – С. 20-29.
2. Круглик, А. С. Чат-бот для людей с ограниченными возможностями здоровья // Молодежный Вестник УГАТУ. – 2022. – №. 1 (26). – С. 42-46.
3. Щеткина, М. А. Приоритеты государственной политики в области социальной защиты на современном этапе // Проблемы управления (Минск). – 2011. – №. 2. – С. 23-30.
4. Постановление Совета Министров Республики Беларусь от 21 декабря 2020 г. № 748 «О Государственной программе «Социальная защита» на 2021–2025 годы [Электронный ресурс] – Режим доступа: <https://mintrud.gov.by/uploads/files/POSTANOVLE\protect\discretionary{\char\hyphenchar\font}{-}{NIE-748-v-red-824.pdf>. – Дата доступа: 01.10.2023.
5. Ясницкий, Л. Н. Интеллектуальные системы // М.: Лаборатория знаний. – 2016. – Т. 221.
6. Технология комплексной поддержки жизненного цикла семантически совместимых интеллектуальных компьютерных систем нового поколения / под общ. ред. В. В. Голенкова. – Минск : Бестпринт, 2023. – 1064 с.
7. НИКА [Электронный ресурс] – Режим доступа: <https://github.com/ostis-apps/nika>. – Дата доступа: 01.10.2023.
8. Садовский, М. Е. Индивидуализация пользовательских интерфейсов интеллектуальных систем на основе семантической модели // Цифровая трансформация. – 2023. – Т. 29. – №. 3. – С. 54-63.

ИСПОЛЬЗОВАНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ РЕШАТЕЛЕЙ ЗАДАЧ В СИСТЕМЕ УПРАВЛЕНИЯ ЖЕЛЕЗНОДОРОЖНОЙ СТАНЦИИ

Ерофеев И. А.

Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ilerofv@gmail.com

Определена актуальность применения автоматизированных систем в управлении железнодорожной станцией. Рассмотрены примеры задач управления. Проанализировано текущее состояние и предложены дальнейшие шаги по разработке интеллектуальной системы управления железнодорожной станцией.

ВВЕДЕНИЕ

Внедрение интеллектуальных систем управления является актуальной задачей для всех видов деятельности человечества. Не исключением является и железнодорожный транспорт, в особенности система управления перевозочным процессом на железной дороге. Важной составляющей системы управления перевозочным процессом на железной дороге является подсистема планирования. Как в любой современной проблеме такого рода, основные сложности возникают в нетипичных случаях, например когда тяжело предсказать значения переменных до начала решения или когда переменные плохо формализуются.

Рассмотрим более детально суть задачи планирования в контексте инфраструктуры железной дороги. На данный момент на железной дороге существуют автоматические планировщики графиков, основанные на классических алгоритмах. Однако их недостатками является малая гибкость и низкая адаптивность. В связи с этим они применяются только в системах долгосрочного и среднесрочного планирования. Касательно же оперативного планирования, для управления поездной и маневровой работой, а так же информационной поддержки процессов на станции, на Белорусской железной дороге существуют две информационно-справочные системы. Это автоматизированная система управления станцией (АСУС) и автоматизированная система управления линейным районом (АСУ ЛР). При этом функции разработки сменно-суточных и текущих планов работы станции реализуются в ручном режиме с последующим вводом информации в базу данных системы, что влияет на их достоверность и задержку появления в системе относительно реального времени[1].

Таким образом одной из наиболее актуальных задач автоматизации на железной дороге на данный момент является внедрение технологий искусственного интеллекта в системы автоматизации этих процессов.

Далее рассмотрим эту идею более подробно на примере задачи формирования оперативных планов.

I. ЗАДАЧА ФОРМИРОВАНИЯ ОПЕРАТИВНЫХ ПЛАНОВ

Для решения задач оперативного управления необходимо собрать информацию, в том числе прогнозную, что само по себе является нетривиальной задачей. Приведём несколько примеров задач, решаемых при разработке оперативных планов:

1. Прогнозирование длительности технологических операций на станции. Например, осмотр вагона с учётом таких признаков как: погода, тип вагона, опыт специалиста и др. **На данный момент** существуют фиксированные нормативы. Таким образом планировщик берёт информацию просто из «таблички».
2. Оптимизация графика приёма составов с учётом длительности тех. процессов, типов и требований составов и т.д.. Минимизация простоя по причине неприятия. **На данный момент** такой график заранее рассчитывается на основании плана подхода поездов. В случаях изменений в графике прибытия поездов, в технологии работы на станции и т.д., разработка оперативных планов ведётся вручную.
3. Динамическое перераспределение нагрузки между станциями, в условиях высокой загрузки одних станций и низкой загрузки других. **На данный момент** перераспределение нагрузки между станциями происходит на основании плана формирования поездов, который рассчитывается один раз в год и не всегда соответствует номинальным параметрам железнодорожной инфраструктуры и объёмам перевозок грузов[2].

На данный момент не существует доступной системы, способной решать эти задачи как на Белорусской железной дороге, так и за рубежом, однако существуют системы, решающие задачи аналогичных классов.

II. АНАЛИЗ СУЩЕСТВУЮЩИХ ПОДХОДОВ

Интеллектуальное прогнозирование длительности операции не является новой задачей,

существуют различные методы решений схожих задач на основе process mining и машинного обучения. В частности поиск проблем и закономерностей в ПО на основании журнала событий. Учитывая существующие на железной дороге правила сбора информации [4] есть все основания предполагать наличие исчерпывающей информации для использования этих методов.

Для модуля оптимизации графика приёма составов можно так же использовать решения машинного обучения [1]. В качестве обучающей выборки и источника входных параметров можно взять упомянутые уже журналы событий и существующую на Белорусской железной дороге информационную систему соответственно.

Задачи динамического перераспределения нагрузки могут быть решены в рамках стандартных алгоритмов. Примером таких решений в программном обеспечении может стать распределение вычислительных ресурсов в вычислительных центрах. Однако более близким нам вариантом будет пример из материального производства, например система динамического распределения складских ресурсов промышленного предприятия [5].

III. ПРЕДЛАГАЕМЫЙ ПОДХОД

Для создания комплексной системы решения задачи формирования оперативных планов нам потребуется обеспечить унификацию и семантическую совместимость компонентов этой системы между собой. Указанный принцип реализуется в рамках Открытой семантической технологии проектирования интеллектуальных систем (*Технологии OSTIS*) [6], которую предлагается положить в основу предлагаемой системы.

Разработка системы в рамках данной технологии подразумевает разработку онтологии предметной области, разработку модулей решателей, управляющего модуля, а так же интеграция указанных элементов в рамках общей мета-системы.

IV. РАЗРАБОТКА ОНТОЛОГИИ

Как было сказано, для решения проблемы совместимости на уровне *Технологии OSTIS* предлагается разработать онтологию единых терминов и понятий в рамках интересующей нас предметной области, которая позволит модулям использовать результаты работы друг-друга. В этой онтологии должны быть описаны все элементы системы, а так же процессы в ней протекающие. За основу предлагается взять онтологии из уже существующих работ и расширить их при необходимости. Примером верхнего уровня сущностной онтологии может быть приведена таблица указанная далее [3]. Однако стоит заметить, что каждый из указанных терминов, сам является

частной предметной областью, которая описывается собственной онтологией.

Таблица 1 – Сущностная онтология верхнего уровня.

Статические объекты (инфраструктура)	Динамические объекты
Железная дорога	Поезд
Отделение	Вагон
Железнодорожная станция	Контейнер
Межгосударственный пункт переезда	Отправка
Район управления ЦУП	Локомотив
Диспетчерский участок	Бригада
Железнодорожное направление	Документ
Железнодорожный участок	
Перегон	
Блок-участок	
Депо	

V. ЗАКЛЮЧЕНИЕ

В работе определена актуальность применения автоматизированных систем в управлении железнодорожной станцией. Рассмотрены примеры задач управления. Проанализировано текущее состояние и предложены дальнейшие шаги по разработке интеллектуальной системы управления железнодорожной станцией.

Дальнейшая разработка предполагает формализацию онтологии средствами OSTIS, реализация соответствующих решателей задач и интеграция разработанных компонентов в единую интеллектуальную систему.

1. Буглак С. В. Развитие информационно-управляющих систем станционного уровня / С.В Буглак // ИТ-ЭС 2020: Инновации. Транспорт. Энергоэффективность. Строительство [Электронный ресурс]: сборник тезисов докладов VI межд. науч.-практ. конф. магистрантов / Белорусский государственный университет транспорта – Гомель: БелГУТ, 2020. – с. 21
2. Информационные технологии на железнодорожном транспорте: учеб. -метод. пособие Ч.2 /Ерофеев А.А., Федоров Е. А. // М-во образования Респ. Беларусь, Белорус. гос. ун-т трансп. БелГУТ – Гомель:БелГУТ, 2015 – 256 с.
3. Интеллектуальная система управления перевозочным процессом на железнодорожном транспорте /А.А Ерофеев // М-во образования Респ. Беларусь, Белорус. гос. ун-т трансп. БелГУТ – Гомель: БелГУТ, 2022 – 407 с.
4. СТП БЧ 15.405-2019 Порядок ведения графика исполненной работы на станциях в условиях автоматизированных систем: Стандарт государственного объединения «Белорусская железная дорога»
5. Плонский В. Ю. Система Динамического перераспределения складских ресурсов промышленного предприятия / Плонский В. Ю., Чистякова Т. Б // Вестник АГТУ. Серия: Управление, вычислительная техника и информатика. – 2020. – №4
6. Открытая технология онтологического проектирования, производства и эксплуатации семантически совместимых гибридных интеллектуальных компьютерных систем / В. В. Голенков, Н. А. Гулякина, Д. В. Шункевич. – Минск : Бестпринт, 2021. – 690 с.

КОЛИЧЕСТВЕННЫЕ ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ ОПЕРАЦИЙ НАД ОБЩЕЙ СЕМАНТИЧЕСКОЙ ПАМЯТЬЮ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Зотов Н. В.

Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: n.zotov@bsuir.by

В данной работе проводится оценка эффективности общей семантической памяти программной платформы для проектирования и реализации интеллектуальных систем. Текущая работа описывает количественную оценку эффективности операций над общей семантической памятью в однопоточной и многопоточной средах.

ВВЕДЕНИЕ

Ранние работы [1, 2] были посвящены описанию модели программной платформы для интеллектуальных систем и её реализации (программной платформе ostis-систем). В процессе модернизации удалось существенно повысить эффективность обработки информации в многопоточной среде. Цель текущей работы заключается в оценке эффективности операций этой программной платформы в однопоточной и многопоточной средах (в частности, операций на общей семантической памяти (sc-памятью)). Актуальность работы обусловлена ростом требований к скорости обработки информации в многопоточных (многопользовательских) средах.

I. ЭФФЕКТИВНОСТЬ ОПЕРАЦИЙ НАД SC-ПАМЯТЬЮ

Текущая *Реализация sc-памяти ostis-платформы* позволяет хранить и представлять *sc-конструкции*, внешние *информационные конструкции*, не принадлежащие *SC-коду*, а также контролировать и согласовать процессы в ней.

Тестирование и оценка эффективности программной платформы ostis-систем были проведены на одной из её последних версий – 0.9.0. На данной версии платформы решена задача управлением процессов в общей семантической памяти. В ходе тестирования были вычислены основные показатели эффективности (производительности) операций над sc-памятью в однопоточной и многопоточной средах: время отклика и пропускная способность [3], а также было вычислено ускорение [4], получаемое при использовании параллелизма при выполнении группы операций одного класса над sc-памятью. В качестве вычислительного устройства использовался ноутбук *HP ProBook Hewlett Packard*, имеющий процессор *Intel(R) Core(TM) i7-4900MQ* (4 ядра по 2 потока) с настроенной частой ядер в 3.20 GHz, RAM размером 16 ГБ и SSD размером 256 ГБ.

Результаты тестирования операций над sc-памятью, включающую реализацию модели управления процессами, показали, что параллель-

ное выполнение операций над sc-памятью эффективно в тех случаях, когда их количество достаточно велико (например, 1,000,000 операций) (Таблица 1).

Таблица 1 – Эффективность использования 4 физических потоков для выполнения 1,000,000 операций над sc-памятью по сравнению с 1 физическим потоком

Количество физических потоков	1 поток		4 потока	
	Время отклика (RT), мс	Время отклика (RT), мс	Ускорение (Speedup), раз	
Операции добавления (изменения)				
Операция создания sc-узла	958,025	369,680	2.591	
Операция создания sc-коннектора	1,299.740	787.001	1.652	
Операция добавления содержимого в файл ostis-системы	29,885.500	9,555.450	3.128	
Операции поиска				
Операция поиска выходящих sc-коннекторов из заданного sc-элемента	642.378	203.005	3.164	
Операция поиска файла ostis-системы по его содержанию	1,608.650	928.555	1.732	
Операции удаления				
Операция удаления sc-элемента	1,850.950	1,746.270	1,060	
Операция удаления выходящих sc-коннекторов из заданного sc-элемента	1,704.620	2,115.500	0.806	

В то же время параллельное выполнение малого числа операций над sc-памятью (например, 100 или 10.000 операций) в ряде случаев может быть хуже их последовательного выполнения (Таблица 2).

Такое поведение связано с особенностями механизмов управления процессов в общей семантической памяти, классами выполняемых операций, а также с их заданными входными значениями в контексте решаемой задачи. Например, все операции поиска sc-конструкций с одними и теми же sc-элементами, выполняемые параллельно, не блокируют друг друга. Например, скорость параллельного выполнения операций над файлами ostis-систем зависит от объёма используемого буфера при чтении внешних информационных конструкций и записи их на диск, а также от длины самих информационных конструкций.

Таблица 2 – Эффективность использования 4 физических потоков для выполнения 100 операций над sc-памятью по сравнению с 1 физическим потоком

Количество физических потоков	4 потока		
	1 поток	4 потока	
	Время отклика (RT), мс	Время отклика (RT), мс	Ускорение (Speedup), раз
Операции добавления (изменения)			
Операция создания sc-узла	0.099	1.306	0.076
Операция создания sc-коннектора	0.150	0.422	0.356
Операция добавления содержимого в файл ostis-системы	9.521	4.128	2.307
Операции поиска			
Операция поиска выходящих sc-коннекторов из заданного sc-элемента	0.530	0.241	2.200
Операция поиска файла ostis-системы по его содержимому	0.339	1.453	0.233
Операции удаления			
Операция удаления sc-элемента	0.144	1.494	0.096
Операция удаления выходящих sc-коннекторов из заданного sc-элемента	0.182	0.938	0.194

II. ЭФФЕКТИВНОСТЬ СЕТЕВЫХ ОПЕРАЦИЙ НАД

Сетевой доступ к sc-памяти обеспечивает серверная подсистема программной платформы ostis-систем, реализованная на основе языков (протоколов) WebSocket и JSON и предоставляющая сетевые операции (команды) над sc-памятью [1].

В процессе тестирования реализации была вычислена пропускная способность её команд. При нагрузочном тестировании использовалась тестовая клиентская система, реализованная на C++. В качестве устройства использовалось то же устройство, что и при тестировании операций над sc-памятью. В результате было выяснено, что при отправке 1000 различных команд: *команд создания sc-элементов, команд обработки содержимого файлов ostis-системы и команд удаления sc-элементов* — время, потраченное на их обработку не превышало 0,2 секунды. При этом в отдельных случаях на обработку 1000 *команд создания sc-элементов* уходило не более 0,14 секунды, *команд удаления sc-элементов* — не более 0,12 секунды, *команд обработки содержимого файлов ostis-системы* — не более 0,10 секунды, *команд поиска sc-конструкций, изоморфных заданному пятиэлементному графу-образцу* — не более 0,45 секунды.

III. ЗАКЛЮЧЕНИЕ

Из результатов тестирования следует, что текущая реализация программной платформы ostis-систем является эффективным средством для обработки распределённой информации как с использованием программного интерфейса, так и сетевого интерфейса и протоколов коммуникации.

1. Zotov, N. Software platform for next-generation intelligent computer systems = Программная платформа для интеллектуальных компьютерных систем нового поколения / N. Zotov // Open Semantic Technologies for Intelligent Systems (OSTIS-2022) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенков [и др.]. – Минск, 2022. – Вып. 6. – С. 297–326.
2. Zotov, N. Design principles, structure, and development prospects of the software platform of ostis-systems = Принципы проектирования, структура и перспективы развития программной платформы ostis-system / N. Zotov // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенков [и др.]. – Минск, 2023. – Вып. 7. – С. 67–76.
3. Бахтизин, В. В. Эффективность программных средств и ее оценка // Доклады Белорусского государственного университета информатики и радиоэлектроники. – 2007. – №. 2 (18). – С. 128-133.
4. Gonnord, L. A Survey on Parallelism and Determinism / L. Gonnord et al. // ACM Computing Surveys. – 2023. – Т. 55. – №. 10. – С. 1-28.

МОДЕЛЬ УПРАВЛЕНИЯ ПРОЦЕССАМИ В ОБЩЕЙ СЕМАНТИЧЕСКОЙ ПАМЯТИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Зотов Н. В.

Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: n.zotov@bsuir.by

В данной работе исследуется задача распределенной обработки информации в общей семантической памяти. Работа представляет новую модель управления процессами, отвечая на растущие требования по обработке информации в многопоточных средах и открытых вычислительных системах.

ВВЕДЕНИЕ

Ранее в работах [1, 2], посвящённых описанию программной платформы для интеллектуальных систем, разрабатываемых по принципам Технологии OSTIS (программной платформе *ostis-систем*), была рассмотрена программная реализация общей семантической памяти (sc-памяти), а также подробно описана реализация её программного интерфейса. Цель текущей работы заключается в решении задачи распределённой обработки информации в общей семантической памяти, функционирующей в многопоточной среде. Актуальность работы обусловлена увеличением объемов обрабатываемой информации, ростом требований к скорости обработки информации и недостаточной производительностью современных открытых вычислительных систем. Новизна работы заключается в предлагаемой модели управления процессами в общей семантической памяти.

I. МОДЕЛЬ SC-ПАМЯТИ В ПЛАТФОРМЕ OSTIS-СИСТЕМ

Программную платформу *ostis-систем* можно рассматривать как систему управления графовыми базами данных (СУБД) [1]. Однако она имеет ряд особенностей по сравнению с современными СУБД [1]:

- обрабатываемые графовые конструкции являются конструкциями SC-кода (sc-конструкциями),
- в её основе заложены событийно-ориентированная обработка информации и основанный на ней многоагентный подход.

Под текущей программной реализацией *sc-памяти* понимается компонент программной платформы *ostis-систем*, осуществляющий хранение sc-конструкций и доступ к ним через соответствующий программный интерфейс [3]. В общем случае *sc-память* выполняет следующие задачи:

- хранение *sc-конструкций*,
- хранение внешних по отношению к SC-коду информационных конструкций (файлов *ostis-системы*),

- доступ (создание, чтение и удаление) к sc-конструкциям, реализуемый через соответствующий программный интерфейс,
- управление процессами (потоками).

В рамках данной программной реализации *ostis-платформы sc-память* представлена в виде набора фиксированного размера N , состоящего из сегментов ($S = \langle s_1, s_2, \dots, s_i, \dots, s_n \rangle$), каждый из которых представляет собой набор фиксированного размера M , состоящей из элементов *sc-памяти* (или ячеек) ($s_i = \langle e_{i1}, e_{i2}, \dots, e_{ij}, \dots, e_{im} \rangle$, $e_{ij} \in E$). Каждый элемент *sc-памяти* e_{ij} (в т.ч. элемент *sc-памяти*, соответствующий *sc-узлу* n_k) содержит синтаксический и семантический классы, уровни доступа соответствующего *sc-элемента*, выраженные в виде бинарных строк $t \in T$ и $r \in R$ соответственно, *sc-адреса элементов sc-памяти*, соответствующих первому входящему и выходящему *sc-коннекторам*, а также число входящих *sc-коннекторов* в заданный *sc-элемент* и число выходящих *sc-коннекторов* из заданного *sc-элемента*. При этом каждый элемент *sc-памяти*, соответствующий *sc-коннектору* c_h дополнительно содержит *sc-адреса элементов sc-памяти*, соответствующих начальному и конечному *sc-элементам* этого *sc-коннектора*, *sc-адреса элементов sc-памяти*, соответствующих предыдущим и следующим входящим и выходящим *sc-коннекторам* ($E = N \cup C$, $n_k \in N$, $c_h \in C$).

Программная реализация *sc-памяти* включает также программную реализацию файловой памяти, которая обеспечивает хранение содержимого L файлов *ostis-систем* F ($F \subset N$, $N \times L \rightarrow F$, $F \times L \rightarrow F$), то есть внешних информационных конструкций (линейных текстов, изображений, видео и т.д.), а также выгрузку *sc-памяти* на дисковое пространство. В качестве структур данных в файловой памяти используются префиксные деревья и линейные списки.

Программный интерфейс реализации *sc-памяти* представляет собой набор методов, позволяющих работать с *sc-памятью*. Его модель можно определить по формуле 1.

$$PI = N^T \times C^{E \times E \times T} \times F^{N \times L} \times \{E^{E \times \{E \cup T\} \times E} \cup E^{\{E \cup T\} \times T \times E} \cup E^{E \times T \times \{E \cup T\}}\} \times T^E \times F^L \times L^F \times \{\top, \perp\}^E \quad (1)$$

II. МОДЕЛЬ УПРАВЛЕНИЯ ПРОЦЕССАМИ В SC-ПАМЯТИ

В реализации *sc-памяти* логически выделяется программная реализация подсистемы управления процессами в ней. Она выполняет следующие задачи:

- создание, чтение и удаление процессов,
- синхронизацию процессов,
- распределение *sc-памяти* между процессами,
- управление событиями.

В *sc-памяти*, обычно, используются потоки процессов. Поток может возникнуть в результате инициированного по событию *sc-агента* или в ответ на запрос к *sc-памяти* через сеть, выполняющих операции чтения и записи в ней.

В предыдущих версиях программной платформы *ostis-систем* для обеспечения согласованного доступа к данным в разделяемой памяти (*sc-памяти*) применялись мьютексы и атомарные операции. В версии 0.9.0 для согласования доступа к структурам данным в *sc-памяти* были реализованы мониторы [3], а также методы для захвата и освобождения ресурсов для потоков-читателей P_r (далее – читателей) и потоков-писателей P_w (далее – писателей) ($P = P_r \cup P_w$).

Монитор $m \in M$ состоит из: счётчика активных читателей, флага, показывающего, активен ли в данный момент какой-либо писатель, очереди читателей и писателей, а также мьютекса, используемого для синхронизации доступа к этим элементам монитора. Очередь читателей и писателей представляет собой последовательность запросов на захват определенного ресурса. Каждый запрос включает в себя уникальный идентификатор потока, тип потока (читатель или писатель) и условную переменную, позволяющую обмениваться сообщениями между процессами (потоками). Эта очередь гарантирует, что ни один поток не останется «голодным». Важно отметить, что захват ресурса методом монитора для чтения одним потоком не блокирует захват этого же ресурса методом монитора для чтения другим потоком.

Распределение *sc-памяти* писателям осуществляется посегментно с использованием специализированной хеш-таблицы T_{ps} , которая позволяет определить, занят ли данный незаполненный сегмент *sc-памяти* другим писателем ($s : S \rightarrow S_{ne}$, $T_{ps} \subseteq P_w \times S_{ne}$). Незаполненный сегмент *sc-памяти* может представлять собой как сегмент с невыделенными ячейками, так и сегмент с освобожденными ячейками. При распределении *sc-памяти* сначала происходит поиск незаполненных

сегментов, которые не используются другими писателями. Если такие сегменты не найдены, то выделяются новые сегменты. Если в *sc-памяти* нет доступного места для новых сегментов, писатели могут использовать сегменты из списка занятых незаполненных сегментов. Для обеспечения согласованного доступа к сегментам для чтения каждый сегмент содержит уникальный монитор ($s_m : S \times M \rightarrow S_m$).

Также кроме сегментов мониторы временно назначаются элементам *sc-памяти* и регистрируемым событиям в ней. Мониторы элементов *sc-памяти* хранятся в специализированной хеш-таблице T_{em} . С помощью данных мониторов синхронизируется доступ к информации о *sc-элементе*, содержащейся в элементе *sc-памяти* ($T_{em} \subset A \times M$). События $v \in V$, их типы $t_v \in T_v$, *sc-агенты*, подписанные на них $A_v \subset A$, а также мониторы этих событий хранятся в единой хеш-таблице T_v . С помощью данных мониторов синхронизируются подписка и отписка на события через единую хеш-таблицу, а также инициирование самих *sc-агентов* ($v_i = \langle t_v^i, A_v^i, m_v^i \rangle$, $t_v^i \in T_v, A_v^i \subset A_v, m_v^i \in M, v_m : V \times M \rightarrow V_m$).

III. ЗАКЛЮЧЕНИЕ

Предложенная модель управления общей семантической памятью позволяет эффективно отслеживать и синхронизировать параллельный доступ к данным. Реализация этой модели демонстрирует значительное увеличение (на 2-3 порядка) пропускной способности параллельного выполнения задач в сравнении с предыдущими версиями платформы. Однако для обеспечения (причинной, последовательной) консистентности процессов и их операций кроме уровня данных необходимо управлять уровнем знаний [4].

1. Zotov, N. Software platform for next-generation intelligent computer systems = Программная платформа для интеллектуальных компьютерных систем нового поколения / N. Zotov // Open Semantic Technologies for Intelligent Systems (OSTIS-2022) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенков [и др.]. – Минск, 2022. – Вып. 6. – С. 297–326.
2. Zotov, N. Implementation of Information Retrieval Subsystem in the Software Platform of ostis-systems = Реализация информационно-поисковой подсистемы в программной платформе *ostis-систем* / N. Zotov // Open Semantic Technologies for Intelligent Systems (OSTIS) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенков [и др.]. – Минск, 2023. – Вып. 7. – С. 77–94.
3. Cole M. I. Algorithmic skeletons: structured management of parallel computation. – London : Pitman, 1989.
4. Miret L. P. Consistency models in modern distributed systems. an approach to eventual consistency // Master. MA thesis. Universitat Politècnica de Valencia, Spain. – 2014.

ИНФОРМАЦИОННЫЕ ХАРАКТЕРИСТИКИ ОНТОЛОГИЧЕСКИХ СТРУКТУР

Ивашенко В. П.

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

В работе рассматриваются онтологические структуры с операционной семантикой на базе графовых динамических структур, рассматриваемых как обобщённые конечные автоматы. Предлагается подход к вычислению информационно-энтропийных характеристик в этих структурах, с этой целью выделяются подвиды этих структур.

ВВЕДЕНИЕ

Способность к решению задач является определяющим качеством интеллектуальных систем [1]. Решение задач, связано с представлением и обработкой знаний, в процессах которых используются графовые структуры: семантические сети [2,3], графы состояний [2], нейронные сети [4,5] и т.д. Семантические сети в первую очередь связаны с денотационной семантикой, тогда как графы состояний в рамках формальных моделей обработки знаний [2] связаны с операционной семантикой, включая игровую [6].

В работах [6,7] рассмотрена взаимосвязь различных видов семантик в языках модели унифицированного семантического представления знаний [2,6,7], использующего графовые структуры [3,5]. Также в работах [5,8] предложен подход к анализу и выявлению признаков (характеристик) графовых структур с целью их сравнения в задачах управления знаниями для систем, управляемых знаниями [3].

С точки зрения управления знаниями для задач сравнения онтологических структур следует выделять семантические сходство и близость [7] для структур различного вида и сложности: отдельных элементов (понятий), статических структур (фрагментов онтологий), динамических структур (фрагментов моделей решения задач). Один из подходов состоит в том [8], что с целью сравнения онтологических структур необходимо построить спецификацию этих структур в рамках модели спецификации знаний [2,7], которая согласуется с метамоделью смыслового пространства [7].

I. АНАЛИЗ И МОДЕЛЬ ПСЕВДОГРАФОВЫХ ДИНАМИЧЕСКИХ СТРУКТУР

В работе [8] рассмотрены два вида структур: (конечные) ациклический граф и сильно связный псевдограф. Произвольная (конечная) псевдографовая структура может быть разбита на свои связные компоненты. Произвольная (конечная) связная псевдографовая структура может иметь различные виды подструктур и, в частности, может быть разбита на два рассмотренных ранее

вида структур: сильносвязные компоненты (подпсевдографы), ациклические графы (подграфы). Однако, в рамках структуры эти подструктуры могут иметь различные отношения и выполнять различные роли. Рассмотрим различные виды подструктур в соответствии с их ролями (отношениями), выполняемыми (имеющимися) в структуре.

Резонаторы – максимальные сильносвязные подпсевдографы.

Датчики (рецепторы) – ациклические подграфы, элементы которых не достижимы ни из одного резонатора.

Раздатчики (эффекторы) – ациклические подграфы, из элементов которых не достижим ни один из резонаторов.

Передатчики (трансммиттеры) – ациклические подграфы, элементы которых достижимы хотя бы из одного резонатора и из элементов которых достижим хотя бы один другой (другой) резонатор.

Раздатчики и передатчики являются потребителями.

Датчики и передатчики являются поставщиками.

Также можно рассматривать порождающие резонаторы (не достижимые из других резонаторов) и потребляющие резонаторы (достижимые из других резонаторов).

Среди резонаторов можно выделить унимодальные (гармонические) резонаторы и многомодальные (негармонические) резонаторы (многомодальная форма волны). Все многомодальные резонаторы являются потребляющими. Каждая мода сдвинута по фазе менее чем на период.

Из любой структуры можно выделить безрезонаторный подграф, который является множеством всех вершин и рёбер, достижимых из элементов рецепторов.

Оставшиеся рёбра совместно с их начальными и конечными вершинами образуют резонаторный подпсевдограф.

Элемент датчика называется элементом-датчиком, если у него отсутствует поставщик.

Элемент раздатчика называется элементом-раздатчиком, если у него отсутствует потребитель.

Если отсутствуют поставщики (датчики или передатчики), элементы которых являются поставщиками порождающему резонатору, то все его элементы являются элементами-датчиками.

Если отсутствуют потребители (раздатчики или передатчики), элементы которых являются потребителями резонатора, то все его элементы являются элементами-раздатчиками.

II. МЕТОД ОПРЕДЕЛЕНИЯ ЁМКОСТНЫХ ХАРАКТЕРИСТИК МОДЕЛЕЙ РЕШЕНИЯ ЗАДАЧ

Метод определения ёмкостных характеристик структур, используемых при решении задач, сводится к следующим принципам.

Каждый резонатор обладает периодом, который является частным НОД T собственного периода и периодов всех (его) поставщиков, от которых он потребляет, к НОД этого НОД (T) и всех его делителей k , для которых свёртка фаз $k * n$ с фазами (с учётом их сдвигов) всех мод (формы волны) сигнала равна $k^{T-k} * ((T - k)!)$. Для каждого элемента потребителя (раздатчика или передатчика) период рассчитывается аналогично, за исключением того, что вместо НОД собственно периода и периодов всех (его) поставщиков, от которых он потребляет, берётся НОД периодов всех его поставщиков.

Для каждого элемента поставщика можно аналогичным образом определить потенциальный период, который является периодом потребителя в инвертированном псевдографе (псевдограф обратного отношения).

Для элементов-датчиков может быть принято, если не принято иначе, что их период равен потенциальному периоду.

Для каждого элемента датчика (поставщика) можно рассчитать период, который равен НОК периодов всех его потребителей (с учётом сдвига по фазе).

Энтропия множества резонаторов рассчитывается на периоде равном НОК периодов всех резонаторов этого множества.

Если резонаторы присутствуют, то (максимально возможная) энтропия всего псевдографа действительна и может быть вычислена как наибольшая энтропия из наибольших энтропий наименьших множеств резонаторов, разрезающих множество (путей) из множества наименьших множеств путей, соединяющих все элементы-датчики со всеми элементами-раздатчиками.

Если все элементы-датчики и элементы-раздатчики являются элементами резонаторов, то (минимально необходимая) энтропия всего псевдографа действительна и может быть вычислена как наименьшая энтропия из наименьших энтропий наименьших множеств резонаторов, разрезающих множество (путей) из множества наимень-

ших множеств путей, соединяющих все элементы-датчики со всеми элементами-раздатчиками.

Если все элементы-датчики и элементы-раздатчики не являются элементами резонаторов, то (минимально необходимая) энтропия всего псевдографа недействительна (мнимая) и вычисляется на периоде равном НОК всех периодов (элементов) подпсевдографов псевдографа и максимальной из длин (простых) путей от элемента-датчика до элемента-раздатчика.

Энтропия несвязного псевдографа может быть вычислена как среднее (минимальное, максимальное и т.п.) энтропий его компонент на периоде равном НОК периодов (на которых производились вычисления для) его компонент.

ЗАКЛЮЧЕНИЕ

Проанализированы конечные динамические структуры, используемые при моделировании решения задач. Выявлены виды подструктур, рассмотренные в рамках соответствующей модели. Перечислены основные принципы метода определения ёмкостных характеристик онтологических структур в моделях решения задач. Полученные результаты ориентированы на применение в задачах выявления сходства онтологических структур с целью управления знаниями в интеллектуальных системах.

1. What is Artificial Intelligence (AI)? Режим доступа: <https://www.ibm.com/topics/artificial-intelligence>. Дата доступа: 13.10.2023.
2. Ивашенко, В. П. Модели решения задач в интеллектуальных системах. В 2 ч. Ч. 1 : Формальные модели обработки информации и параллельные модели решения задач : учеб. метод. пособие / В. П. Ивашенко. — Минск : БГУИР, 2020. — 79 с.
3. The standardization of intelligent computer systems as a key challenge of the current stage of development of artificial intelligence technologies / V. V. Golenkov [et al.] // Open Semantic Technologies for Intelligent Systems (OSTIS-2020) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голеньков (гл. ред.) [и др.]. — Минск, 2020. — Вып. 4. — С. 73–88.
4. Нейросетевые технологии обработки данных : учеб. пособие / В. А. Головкин, В. В. Краснопрошин. — Минск : БГУ, 2017. — 263 с.
5. Grattarola, D. Livi, L. Alippi, C. Learning Graph Cellular Automata / D. Grattarola, L. Livi, C. Alippi, 2021, 2110.14237, arXiv.
6. Ивашенко, В. П. Операционная семантика многоагентных систем обработки знаний. / В. П. Ивашенко // Информационные технологии и системы 2020 (ИТС 2020). — Минск: БГУИР, 2020. — С. 78–79.
7. Ivashenko, V. Semantic space integration of logical knowledge representation and knowledge processing models = Интеграция логических моделей представления и обработки знаний в смысловом пространстве / V. Ivashenko // Open Semantic Technologies for Intelligent Systems (OSTIS) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голеньков [и др.]. — Минск, 2023. — Вып. 7. — С. 95–114.
8. Ivashenko, V. Structures and Measures in Knowledge Processing Models / V. Ivashenko // PRIP'2023 proceedings. — Minsk, BSU, 2023. — 16–21 pp.

ОНТОЛОГИЧЕСКИЕ СТРУКТУРЫ И ПАРАМЕТРИЗОВАННЫЕ МНОГОЗНАЧНЫЕ ЛОГИКИ

Ивашенко В. П.

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

Рассматриваются подход к параметризованному представлению онтологических структур, на основе пространства их образов и их признаков и в параметризованных нечётких логических моделях.

ВВЕДЕНИЕ

$$\tau_i \in \{\min, *, \tau_L\}$$

В интеллектуальных системах при обработке знаний одним из важных принципов является принцип учёта НЕ- факторов знания. Нечёткость является одним из НЕ- факторов знания. Нечёткая логика относится к многозначным логикам. Многозначные логики, так же могут использоваться для выражения других НЕ- факторов: неопределённости (логика неопределённости Клини), неточности (Логика Н.А. Васильева и многозначные логики) и т.д.

Для разных видов НЕ- факторов и условий их проявления существуют логики (или частные разновидности логик), ориентированные на поддержку их представления (выражения). Однако, не всегда эти логики (или их разновидности) взаимосвязаны и легко совместимы. Взаимосвязанность логик (и их разновидностей) подразумевает возможность простого, плавного (непрерывного) перехода от одной из них к другой.

С другой стороны существуют математическая параметризация базовых структур и абстракций, используемых для построения этих логик. Эта параметризация может обеспечивать взаимосвязь логик, однако эта взаимосвязь носит зачастую крайне абстрактный характер и её проблематично выразить на предметно- онтологическом уровне.

I. КОНЪЮНКЦИЯ В МНОГОЗНАЧНЫХ ЛОГИКАХ

В нечёткой логике конъюнкция выражается треугольной нормой. Треугольная норма может быть порождена с помощью аддитивного (мультипликативного) генератора из монотонно либо возрастающих, либо убывающих функций с областью определения $[0; 1]$ и областью прибытия $[0; +\infty]$ ($[0; 1]$).

Любая непрерывная треугольная норма может быть выражена строго убывающей биекцией f на интервале $[0; 1]$ и прямой суммой треугольных норм: Гёделя (\min), произведения ($*$), Лукасевича τ_L .

$$\tau_I^f(x, y) =$$

$$f\left(\min_{(\tau_i, a_i, b_i) \in I} \left\{ a_i + (b_i - a_i) * \tau_i \left(\frac{x - a_i}{b_i - a_i}, \frac{y - a_i}{b_i - a_i} \right) \right\}, \min(x, y) \right)$$

Кроме этого существуют параметризованные семейства треугольных норм (семейство Швейцера- Склера и семейство Франка), которые включают три перечисленные треугольные нормы (семейство Швейцера- Склера также включает драстическое произведение).

В общем случае треугольные нормы семейства Швейцера- Склера выражаются:

$$(\max(0, x^p + y^p - 1))^{1/p}$$

В общем случае треугольные нормы семейства Франка выражаются:

$$\log_p \left(1 + \frac{(p^x - 1) * (p^y - 1)}{(p - 1)} \right)$$

Будем рассматривать нечёткость как усреднённую характеристику по экстенсионалу понятия, выражаемого предикатом.

$$P(x) = \lambda y R(y, x)$$

В случае конечного экстенсионала, например:

$$\lambda y R(y, x) = \frac{1}{n} * \sum_{i=1}^n R(y, x)$$

В общем случае среднее не обязательно является средним арифметическим, но задаёт (монотонную) нечёткую меру.

II. ПАРАМЕТРИЗОВАННЫЕ МНОГОЗНАЧНЫЕ ЛОГИКИ

Кроме экстенсионала понятие обладает интенционалом. Рассмотрим пространство образов для соответствующего (двоичного) признакового пространства.

$$P = \{0, 1\}^k$$

Каждый предикат предлагается параметризовать вектором пространства образов.

10	11	01	00	00	00	00	00
10	11	11	11	01	00	00	00
10	10	11	11	01	01	00	00
10	10	10	11	01	01	01	00
10	10	10	10	01	01	01	01
11	11	11	11	00	00	00	00
11	11	00	00	00	00	00	00
11	11	10	10	00	00	00	00

Для параметра треугольных норм семейства Швейцера-Скляра введём преобразование для отображения на интервал $[-1; 1]$:

$$c = \frac{p}{p-2}$$

и обратное преобразование:

$$p = \frac{2 * c}{c-1}$$

Аналогично для параметра треугольных норм семейства Франка введём преобразование для отображения на интервал $[-1; 1]$:

$$g = \frac{1-p}{1+p}$$

и обратное преобразование:

$$p = \frac{g+1}{g-1}$$

Ниже в таблицах приведены примеры для заданных векторов пространства образов, параметр p в таблице 1 определялся из выполнения условия:

$$(p^x - 1) * (p^y - 1) = (p^z - 1) * (p - 1)$$

Таблица 1 – Значение параметров по Франку

x	y	$\tau(x, y)$	p	g	cos	
1/4	1/4	1/8	0.027	0.95	1/2	
1/2	1/2	3/8	0.008	0.99	1/2	1/2
1/2	1/2	1/4	1	0	0	0
1/2	1/2	1/8	131.	-.99	-1/2	-1/2
1/2	1/2	0	$+\infty$	-1	-1	-1
1/2	1/2	1/2	0	1	1	1
1/4	1/4	1/4	0	1	1	1
1/2	1/4	1/4	0	1	1/2	1

В столбце cos вычислен косинус угла между векторами пространства образов, где 1 трактуется как 1, а 0 – как -1.

Таблица 2 – Значение параметров по Швейцеру-Скляру

x	y	$\tau(x, y)$	p	c	cos	
1/4	1/4	1/8	-0.69	0.26	1/2	
1/2	1/2	3/8	-1.9	0.49	1/2	1/2
1/2	1/2	1/4	0	0	0	0
1/2	1/2	1/8	0.69	-.53	-1/2	-1/2
1/2	1/2	0	1	-1	-1	-1
1/2	1/2	1/2	$-\infty$	1	1	1
1/4	1/4	1/4	$-\infty$	1	1	1
1/2	1/4	1/4	$-\infty$	1	1/2	1

В параметризованной нечёткой логике нечёткие логические n -арные операции принадлежат ко множеству:

$$([1; 0] \times P)^{([1; 0] \times P)^n}$$

Каждый аргумент кроме нечёткой степени истинности содержит вектор пространства образов.

Каждая операция φ_λ^P вычисляет кроме степени истинности вектор пространства образов с помощью $\psi_P(x_2, y_2)$.

$$\varphi_\lambda^P(x, y) = \langle \varphi_{\lambda(x_2, y_2)}(x_1, y_1), \psi_P(x_2, y_2) \rangle,$$

где $\varphi_{\lambda(x_2, y_2)}$ – нечёткая операция из параметризованного семейства нечётких операций, а ψ_P – операция вычисляющая вектор пространства образов. Для конструктивного определения параметризованной логики необходимо определить функции над векторами $\varphi_{\lambda(x_2, y_2)}$ и ψ_P .

III. ЗАКЛЮЧЕНИЕ

Введено понятие операций параметризованной нечёткой логики, с целью интерпретации операций и выражений нечёткой логики на онтологических моделях, для которых задаётся пространство образов и соответствующее признаковое пространство.

1. What is Artificial Intelligence (AI)? Режим доступа: <https://www.ibm.com/topics/artificial-intelligence> Дата доступа: 13.10.2023.
2. Ивашенко, В. П. Модели решения задач в интеллектуальных системах. В 2 ч. Ч. 1 : Формальные модели обработки информации и параллельные модели решения задач : учеб. метод. пособие / В. П. Ивашенко. – Минск : БГУИР, 2020. – 79 с.
3. The standardization of intelligent computer systems as a key challenge of the current stage of development of artificial intelligence technologies / V. V. Golenkov [et al.] // Open Semantic Technologies for Intelligent Systems (OSTIS-2020) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенько (гл. ред.) [и др.]. – Минск, 2020. – Вып. 4. – С. 73–88.
4. Нейросетевые технологии обработки данных : учеб. пособие / В. А. Головкин, В. В. Краснопрошин. – Минск : БГУ, 2017. – 263 с.
5. Grattarola, D. Livi, L. Alippi, C. Learning Graph Cellular Automata / D. Grattarola, L. Livi, C. Alippi, 2021, 2110.14237, arXiv.
6. Ивашенко, В. П. Операционная семантика многоагентных систем обработки знаний. / В. П. Ивашенко // Информационные технологии и системы 2020 (ИТС 2020). – Минск: БГУИР, 2020. – С. 78–79.
7. Ivashenko, V. Semantic space integration of logical knowledge representation and knowledge processing models = Интеграция логических моделей представления и обработки знаний в смысловом пространстве / V. Ivashenko // Open Semantic Technologies for Intelligent Systems (OSTIS) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенько [и др.]. – Минск, 2023. – Вып. 7. – С. 95–114.
8. Ivashenko, V. Structures and Measures in Knowledge Processing Models / V. Ivashenko // PRIP'2023 proceedings. – Minsk, BSU, 2023. (in press)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ РАЗЛИЧНЫХ ПОДХОДОВ К ПРОЕКТИРОВАНИЮ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ

Казаченко Е. А.

Кафедра интеллектуальных информационных технологии,
Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: yauheni.kazachenka@gmail.com

В данной работе рассматривается сравнительный анализ различных подходов к проектированию пользовательских интерфейсов. На основании проведенного анализа сделан вывод об целесообразности использования компонентного подхода при построении пользовательских интерфейсов.

ВВЕДЕНИЕ

На сегодняшний день существует огромное количество систем, которые выполняют определенные задачи, в зависимости от требований предъявляемой к предметной области. И у каждого из этих систем реализованы различные методы взаимодействия между пользователем и системой. Методы взаимодействия между пользователем и системой происходят, как правило через **интерфейс пользователя**.

Эффективность использования программной системы напрямую зависит от ее пользовательского интерфейса (UI). Однако, несмотря на то, что интерфейсы развиваются, совершенствуются, они имеют целый ряд проблем, связанных с трудоемкостью их разработки и процессом их сопровождения. Трудоемкость разработки UI обусловлена не столько сложностью разработки пользовательского интерфейса, сколько отсутствием хорошо продуманных технологий в процессе проектирования пользовательских интерфейсов [1]. Поэтому важно правильно подобрать разработчикам правильный подход к проектированию пользовательских интерфейсов, поскольку это может существенно повлиять на эффективность разработки пользовательских интерфейсов, повысить качество проектирования пользовательских интерфейсов, снизить требования к квалификации разработчика и пользователя, а также упростить дальнейшую интеграцию компонентов в другие системы. Правильный выбор подхода к проектированию UI может существенно повлиять на удобство использования и удовлетворение пользователя, а также на успех продукта на рынке. С течением времени, с развитием технологий и появлением новых методологий, появилось множество подходов к проектированию пользовательских интерфейсов.

Целью работы является проведение сравнительного анализа между различными подходами к проектированию пользовательского интерфейсов с целью выявить преимущества, недостатки и области применения различных подходов проектирования.

I. ОБЗОР ПОДХОДОВ К ПРОЕКТИРОВАНИЮ ПОЛЬЗОВАТЕЛЬСКИХ ИНТЕРФЕЙСОВ

Проектирование пользовательских интерфейсов (ПИ) является важной составляющей разработки программного обеспечения, поскольку ПИ имеет прямое влияние на удобство использования продукта, а также влияет на эффективность разработки системы. На сегодняшний день существуют множество подходов к проектированию ПИ, каждый из которых предоставляет уникальные методы и инструменты для создания интерфейсов, которые соответствуют потребностям и ожиданиям пользователей. Однако, как и в любой области, проектирование пользовательских интерфейсов имеет свои особенности и вызовы.

Один из подходов к проектированию пользовательских интерфейсов — традиционный подход, который основан на последовательном выполнении этапов разработки, начиная с анализа требований и заканчивая внедрением и поддержкой системы. Данный подход к проектированию подразумевает строгий контроль над временем и бюджетом проекта, а также акцент на формальной документации и разделении ролей между дизайнерами, разработчиками и тестировщиками [2]. Благодаря этому обеспечивается четкое определение требований и целей, что позволяет снизить риск ошибок при проектировании ПИ. Кроме этого подход позволяет создать детальные макеты и дизайн пользовательского интерфейса, что способствует точному воплощению концепций.

Как правило, данный подход используется в проектах, где необходимо соблюдение строгих стандартов и предсказуемость работы интерфейса, например, в области медицинского оборудования и авиационных систем, где традиционный подход обеспечивает более формализованный и структурированный процесс разработки. Недостатком традиционного подхода проектирования ПИ является долгий цикл разработки, что может привести к задержкам в выпуске продукта. Подход не всегда подходит для проектов, где требуется быстрое реагирование на изменяющиеся требования или рыночные условия. Кроме этого он может не всегда полностью учесть пользова-

тельные потребности, так как изменения могут быть сложными и дорогостоящими после начала разработки.

Другим подходом к проектированию пользовательского интерфейса является объектно-ориентированный подход, который представляет собой методологию, основанная на концепции объектов и их взаимодействия. При таком подходе ПИ разрабатывается с использованием концепции объектов, которые объединяют в себе все данные и функциональность. Основная идея этого подхода заключается в том, чтобы абстрагировать интерфейс как объект, позволяя ему взаимодействовать и создавать иерархии и отношения. Каждый объект представляет собой определенный элемент интерфейса и инкапсулирует данные и методы, необходимые для его функционирования. Объекты могут взаимодействовать между собой через определенные интерфейсы, что способствует модульности и повторному использованию кода. Такой подход облегчает анализ и моделирование пользовательских сценариев и позволяет более гибко адаптироваться к изменяющимся потребностям пользователей.

Тем не менее, следует учитывать и недостатки объектно-ориентированного подхода. Сложность проектирования и разработки в объектно-ориентированных системах может быть высокой, требуя от разработчиков глубокого понимания объектно-ориентированных концепций. Кроме того, несоблюдение принципов правильного проектирования может привести к чрезмерной сложности и избыточности кода. Наконец, объектно-ориентированный подход не всегда является оптимальным для маленьких и простых интерфейсов, где издержки на его применение могут превышать полученные преимущества.

Одним из наиболее распространенных подходов к проектированию ПИ является компонентное проектирование, который представляет собой подход, который основан на идее разделения системы на независимые компоненты, представляющие собой логически связанные модули, которые могут быть повторно использованы для создания пользовательских интерфейсов различного вида сложности. Разделение программы на компоненты обеспечивает надежность, производительность и адаптивность продукта [3]. Одним из главных преимуществ использования компонентов является их возможность повторного использования. Компонент, созданный для одной части приложения, может быть использован в другой части приложения, что сокращает объем кода и ускоряет процесс разработки. Кроме того, использование компонентов позволяет легко изменять и дополнять функциональность приложения без необходимости изменять весь код [4]. Компоненты также улучшают структуру приложения, делая его более понятным и удобным для обслужи-

вания и модификации в будущем. Как правило, компонентный подход к проектированию пользовательского интерфейса широко используется в веб-разработке, где существует большое количество готовых компонентов, которые можно использовать для быстрой разработки интерфейса. Однако, несмотря на все вышеперечисленные преимущества компонентное проектирование обладает также своими недостатками, например, использование данного подхода может потребовать дополнительных затрат на обучение разработчиков, также по мере увеличения числа компонентов возникают проблемы с обслуживанием, тестированием и отладкой пользовательского интерфейса.

II. ЗАКЛЮЧЕНИЕ

Основываясь на сравнительном анализе между различными подходами к проектированию пользовательских интерфейсов, можно сделать вывод, что проектирование ПИ основанный на компонентах, имеет ряд преимуществ перед остальными подходами. Так компонентное проектирование обеспечивает большую возможность повторного использования, гибкость и масштабируемость компонентов пользовательского интерфейса в отличие от других подходов, что в конечном итоге может сэкономить время и усилия в процессе разработки. Кроме того, компонентный дизайн способствует более модульному и организованному подходу к дизайну пользовательского интерфейса, что может привести к повышению удобства обслуживания и упрощению совместной работы членов команды.

Однако выбор между подходами проектированию ПИ напрямую зависят от конкретных требований проекта, имеющихся навыков и ресурсов, а также предпочтений команды разработчиков.

1. Казаченко, Е. А. Использование компонентного подхода при проектировании пользовательских интерфейсов / Е. А. Казаченко // Информационные технологии и управление : материалы 59-ой научной конференции аспирантов, магистрантов и студентов, Минск, 17–21 апреля 2023 года / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. — Минск, 2023. — С. 73–74.
2. Научные статьи.РУ [Электронный ресурс] / От простого к сложному: разбираемся в подходах к созданию информационных систем. — Режим доступа: <https://nauchniestati.ru/spravka/podhodyk-sozdaniyu-informacionnyh-sistem/>. — Дата доступа: 15.10.2023.
3. Голенков, В. В. Интеллектуальный пользовательский интерфейс : учеб.-метод. пособие / В. В. Голенков, Н. А. Гулякина, Д. Г. Колб. — Минск : БГУИР, 2013. — 66 с.
4. Benefits of Componentization [Electronic resource] // Windows Embedded Standard, 2009 — Mode of access: <https://learn.microsoft.com/en-us/previous-versions/windows/embedded/bb521311>. — Date of access: 15.10.2023.

ПОДХОДЫ К АНАЛИЗУ БИОМЕТРИЧЕСКИХ ДАННЫХ В ИНТЕЛЛЕКТУАЛЬНЫХ МЕДИЦИНСКИХ СИСТЕМАХ

Крищенко В. А., Сальников Д. А., Загорский А. Г., Захарьев В. А.
Кафедра интеллектуальных информационных технологий, Кафедра систем управления,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {krish, d.salnikov, a.zagorskiy, zahariev}@bsuir.by

Доклад посвящен анализу предметной области связанной с построением интеллектуальных медицинских систем на основе анализа биометрических характеристик пользователя, а именно, на основе вычисления и анализа электродермальной активности.

ВВЕДЕНИЕ

В настоящее время, в связи с широким распространением персональной вычислительной техники и применения ее в различных областях жизни, все чаще используются методы биометрического распознавания, основанные на вычислении и анализе индивидуальных характеристик пользователя. Это приводит к стремительному росту рынка биометрических технологий. Уже сейчас технологии биометрического распознавания становятся удобным инструментом для решения широкого круга задач.

I. БИОМЕТРИЧЕСКИЕ ДАННЫЕ В ИНТЕЛЛЕКТУАЛЬНЫХ МЕДИЦИНСКИХ СИСТЕМАХ

Выделяют два типа систем биометрических данных:

- статические биометрические данные, т.е. уникальные признаки, полученные человеком от рождения (например: ДНК, отпечатки пальцев, радужная оболочка и др.);
- динамические данные – характеристики, приобретенные со временем или способные меняться с возрастом или под внешним воздействием (например: голос, динамика воспроизведения подписи, походка и др.) [1].

На мировом рынке биометрических систем активно применяются технологии, основанные на распознавании и использовании следующих биометрических данных:

1. отпечатки пальцев (составляют более 50,0% всего объема рынка);
2. изображение лица (21,6%);
3. изображение радужной оболочки глаза (10,2%);
4. голос (4,0%);
5. рисунок вен (3,0%);
6. геометрия ладони, ДНК и иное (7,0%) [2].

Так же, стоит отметить, что наблюдается переход от использования биометрических технологий в обычных системах к интеллектуальным системам.

В современных медицинских системах так же широко используются различные биометрические данные, такие как изменения сердечного ритма, уровень кислорода в крови, активность

мозга, тепловое излучение тела, давление и другие параметры физиологической активности. Эти данные могут быть получены с помощью носимых устройств, мониторов здоровья или специальных медицинских приборов, и используются для диагностики, лечения и мониторинга различных состояний. Одним из таких показателей является электродермальная активность (ЭДА). Электродермальная активность относится к динамическим биометрическим показателям, обозначающим изменения электрической активности кожи. ЭДА отражает изменения в электрической активности кожи, которые коррелируют с уровнем стресса и уровнем проводимости кожи (УПК). Достоинством данного показателя является его применимость и доступность измерений.

На рис. 1 показан типичный сигнал ЭДА, снятый с запястья. Сигнал ЭДА раскладывается по двум физическим параметрам: реакция проводимости кожи (РПК) – динамическая компонента реакций и уровень проводимости кожи (УПК) – более интегрированный показатель (более медленное изменение во времени) [3].

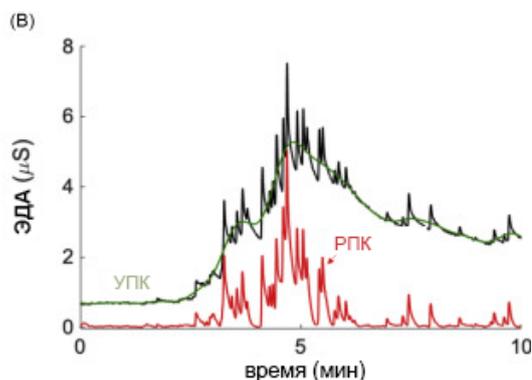


Рис. 1 – Типичный сигнал ЭДА

Спектральное содержание ЭДА в основном ограничивается диапазоном 0,0045-0,15 Гц и поэтому для анализа интересующих откликов достаточно относительно низкой частоты дискретизации (например, 4-20 Гц). Амплитуда РПК измеренная на запястье, как правило, ниже, чем на обычных участках, такие как ладонь или пальцы, и может составлять от нескольких десятков наносекунд до нескольких микросекунд. Ампли-

туда базовой линии ЭДА сильно варьируется как внутри человека, так и между людьми и может достигать 0,1 мкс при типичной записи на запястье, поднимаясь в определенных условиях до десятков мкс или даже до сотен мкс. [3]

ЭДА успешно используется для контроля за состоянием человека при выполнении разных видов деятельности, в частности диагностика функционального состояния.

II. ПЕРСПЕКТИВЫ РАЗВИТИЯ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ

Есть несколько потенциальных областей применения относительно перспектив развития ЭДА в интеллектуальных системах [2]:

1. Биометрическая идентификация.
2. Оценка эмоционального состояния.
3. Приложения в психологии и медицине.
4. Улучшение интерфейсов человек-машина.

Показатель ЭДА может использоваться в качестве одной из ведущих физически измеряемых величин при построении "Интеллектуальной системы персонального медицинского обслуживания который представляет собой OSTIS-систему. [4]

OSTIS-система должна состоять из базы знаний и агентов [5]. Любая OSTIS-система кроме непосредственно семантического уровня должна включать подсистему обработки и анализа сигналов. Для согласования обработки информации в данных подсистемах необходимо произвести формализацию базовых понятий, необходимых для построения предметной области и онтологии, относительно прикладной области для которой будет разрабатываться система. Будет приведена декомпозиция области сбора и хранения биометрических данных. Пример фрагмента базы знаний представлен на рис. 2.

III. ЗАКЛЮЧЕНИЕ

Таким образом, можно сделать вывод, что методы биометрического распознавания основанные на обработке электродермальной активности являются перспективными подходом к анализу биометрических данных.

ЭДА имеет большой потенциал в интеллектуальных системах, и ее применение может значительно расширяться в будущем. Так, недавние исследования в области методологии измерений и интерпретации данных продвинули возможности применения ЭДА как в лаборатории, так и для портативных устройств. Хотя, стандарты высококачественной записи трудно обеспечить при амбулаторных измерениях, недавние и продолжающиеся разработки в области электродов и обработки сигналов могут сократить этот разрыв. Однако, на данный момент еще требуется дальнейшее исследование и разработка технологий для оптимального использования ЭДА в различных областях.

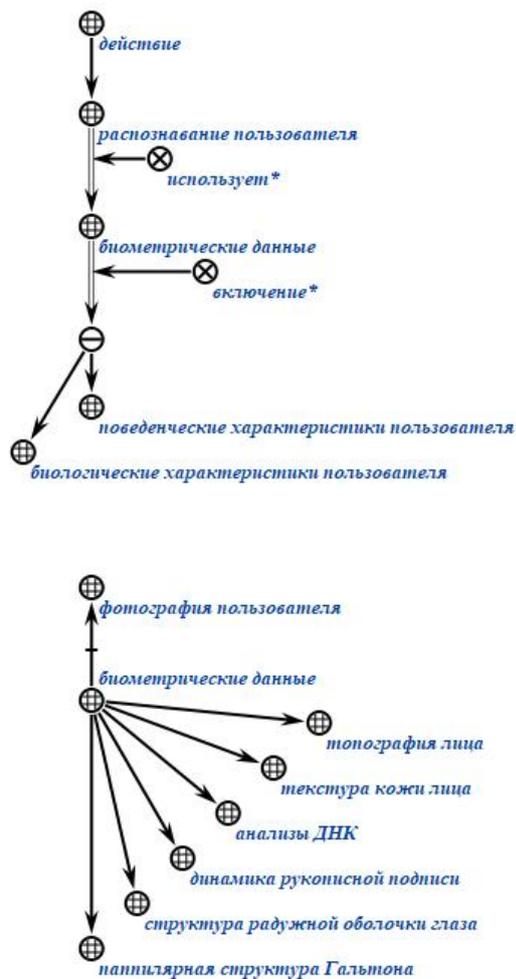


Рис. 2 – Пример формализации понятия биометрических данных в SCg-коде

СПИСОК ЛИТЕРАТУРЫ

1. Смирнова, Я. В. Особенности содержания понятия «биометрические данные» в праве Европейского Союза / Я. В. Смирнова // Вестник Университета имени О.Е. Кутафина. – 2021. – № 7. – С. 160–167.
2. Куликов А. А. Применение биометрических систем в технологиях идентификации лиц. / А. А. Куликов // Russian Technological Journal. – 2021. № 9(3). – С. 7–14.
3. Hugo F. Posada-Quintero Innovations in electrodermal activity data collection and signal processing: A systematic review / Hugo F. Posada-Quintero, Ki H Chon // USA: MDPI: Sensors. – 2020. – Vol. 20, № 2. – P. 479
4. Сальников, Д. А. Модели и средства проектирования интеллектуальных систем персонального медицинского обслуживания / Д. А. Сальников [и др.] // Информационные технологии и системы 2023 (ИТС 2023) = Information Technologies and Systems 2023 (ITS 2023) : материалы международной научной конференции, Минск, 22 ноября 2023 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол.: Л. Ю. Шилин [и др.]. — Минск, 2023.
5. Голенков, В. В. Открытая технология онтологического проектирования, производства и эксплуатации семантически совместимых гибридных интеллектуальных компьютерных систем / В. В. Голенков, Н. А. Гулякина, Д. В. Шункевич. – Минск: Бестпринт, 2021. – 690 с.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АРХИТЕКТУР RNN ДЛЯ РЕШЕНИЯ ЗАДАЧ КЛАССИФИКАЦИИ ПРИ ОБРАБОТКЕ ВИДЕО

Курочка К. С., Башаримов Ю. С.

Кафедра информационных технологий факультета автоматизированных и информационных систем
Гомельского государственного технического университета имени П.О. Сухого

Гомель, Республика Беларусь

E-mail: kurochka@gstu.by, basharymauyury@gmail.com

Выбор архитектуры нейронной сети для решения задачи классификации и распознавания образов является весьма важным этапом, влияющим как на скорость обучения сети, так и на точность и уровень ошибок полученной модели. В данной работе проводится сравнительный анализ различных архитектур рекуррентных нейронных сетей на примере решения задачи классификации объектов в видеопотоке.

ВВЕДЕНИЕ

Классификация видео является сложной задачей, требующей обработки большого количества данных и выявления скрытых закономерностей. В последние годы нейросети стали неотъемлемой частью инструментария для решения задач классификации видео. Архитектуры нейросетей, такие как Convolutional Neural Networks (CNN) и Recurrent Neural Networks (RNN) проявили себя как мощные инструменты для обработки видеоданных и выявления их характеристик [1, 2].

I. ОБЗОР ОСНОВНЫХ АРХИТЕКТУР

В последние годы было разработано множество архитектур нейросетей, специально предназначенных для классификации видео [1]. CNN – это популярная архитектура нейросети для классификации видео. Она извлекает пространственные признаки из кадров видео и моделируют их зависимости. Для классификации видео с помощью CNN сначала нужно подготовить данные для обучения и тестирования, а затем выбрать или создать подходящую архитектуру CNN для видео. После этого нужно обучить модель на тренировочных данных, минимизируя функцию потерь, и протестировать модель на тестовых данных, вычисляя метрики оценки. В результате модель сможет предсказывать метки классов действий для тестовых фрагментов видео.

LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit) – это два типа (RNN) (см. рис. 1), которые хорошо подходят для моделирования последовательных данных, таких как текст или речь. Они способны запоминать долгосрочные зависимости между элементами последовательности и избегать проблемы затухания или взрыва градиента. Эти архитектуры также могут быть адаптированы для классификации картинок и видео.

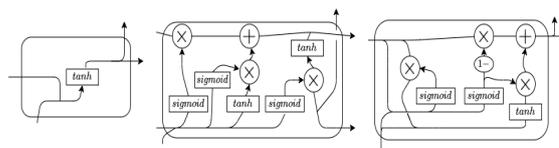


Рис. 1 – Слева на право RNN LSTM GRU

Для обработки видеоданных можно использовать LSTM или GRU по-разному [3]: либо к выходам других нейросетей, извлекающих пространственные признаки, либо к пикселям или патчам напрямую. Это позволяет учитывать как пространственную, так и временную информацию. Но эти подходы имеют недостатки: вычислительную сложность, большой объем данных и сложность интерпретации [4]. Поэтому нужно выбирать архитектуру нейросети в зависимости от задачи.

Для улучшения производительности и точности классификации видео, часто применяются комбинированные архитектуры, которые объединяют свойства CNN и RNN. Например, архитектура Convolutional Recurrent Neural Network (CRNN) [5] (см. рис. 2) комбинирует сверточные слои CNN для извлечения пространственных признаков с рекуррентными слоями RNN для моделирования временных зависимостей в видео.



Рис. 2 – Общая архитектура модели CNN RNN

Такой подход заключается в одновременном выполнении сегментации и классификации. Это называется моделью CNN RNN, специально разработанной для задач прогнозирования последовательности с пространственными входными данными, такими как изображения или видео. Эта архитектура включает в себя использование слоев CNN для извлечения признаков на входных данных в сочетании с RNN для выполнения предсказания последовательности векторов признаков.

Сверточный автоэнкодер сжимает и восстанавливает изображения или видео. Его можно

использовать вместо CNN в CRNN, подавая на вход скрытое представление изображения или видео. Это дает более абстрактные и семантические признаки, но требует больше усилий для обучения и стабилизации. Модель CNN RNN сочетает компьютерное зрение и обработку естественного языка и решает сложные задачи, такие как преобразование видео.

3D-архитектуры нейросетей представляют собой расширение CNN или RNN для работы с трехмерными пространственными и временными данными видео [6]. Они позволяют учитывать изменения во времени и пространстве одновременно, что может быть полезным для классификации видео (см. рис. 3).

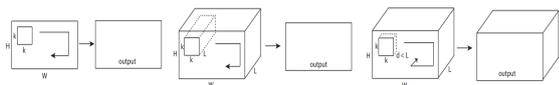


Рис. 3 – Слева на право 2D CNN, 2D CNN с 3D данными и 3D CNN

Архитектура представляет видео как объемные данные и расширяет CNN до третьего, временного измерения заменяя 2D-фильтры 3D-эквивалентами. В отличие от 2D CNN, 3D CNN принимают последовательность видеок кадров в качестве входных данных. 3D-архитектуры нейросетей имеют преимущества перед CNN в классификации видео. Они учитывают пространственно-временные зависимости, изучают временные паттерны, улучшают точность классификации и могут быть компактными. Однако требуют больше ресурсов и данных, и их настройка сложнее.

II. СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Оценим производительность архитектур нейронных сетей на тестовом наборе данных Kinetics (табл. 1) [7, 8].

Таблица 1 – Сравнение разных архитектур нейросетей

Архитектура	Набор-данных	Точность (%)	F1-мера (%)	Скорость (кадров/сек)
RNN	Kinetics	51.01	50.9	15.1
2D CNN	Kinetics	48.91	48.8	30.1
2D CNN с 3D данными	Kinetics	46.11	46.0	25.1
3D CNN (C3D)	Kinetics	51.62	51.5	10.2
3D CNN (I3D)	Kinetics	71.13	71.0	53
CRNN	Kinetics	51.01	50.9	47.1

III. ЗАКЛЮЧЕНИЕ

В завершение, выбор между 3D и 2D нейросетями зависит от задачи, данных и ресурсов. Для распознавания действий в видео лучше 3D CNN или CRNN, которые учитывают пространственные и временные зависимости. Для маленьких данных лучше классические архитектуры. 3D-архитектуры требуют больше ресурсов, чем CRNN. Для классификации видеопотока предлагается CRNN с автоэнкодером вместо CNN. Это ускоряет обучение и повышает качество. Параметры модели: размер изображения – 64 x 64, размер видео – 16 кадров, размер скрытого слоя – 128, размер mini-batch – 32, скорость обучения – 0.001, количество эпох – 50, функция потерь – перекрестная энтропия, оптимизатор – Adam.

СПИСОК ЛИТЕРАТУРЫ

1. Z. Zou, Z. Shi, Y. Guo, and J. Ye Object Detection in 20 Years: A Survey // March 2023 Proceedings of the IEEE. – IEEE, 2023. – С. 257–276.
2. Kurachka K. S., Tsalka I. M. Vertebrae detection in X-ray images based on deep convolutional neural networks // 2017 IEEE 14th International Scientific Conference on Informatics. – IEEE, 2017. – С. 194–196.
3. Олисеенко, В. Д., Абрамов, М. В., Тулупьев, А. Л. Нейронные сети lstm и gru в приложении к задаче многоклассовой классификации текстовых постов пользователей социальных сетей // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – СПб, 2021. – С. 130–141.
4. Kurochka K., Panarin K. Algorithm for real-time binary classification of adenomas and norms images obtained by confocal microscopy // 15th International Conference Mechatronic Systems and Materials, MSM 2020. – 2020. – С. 9202107–9202107.
5. Donahue J, Hendricks LA, Rohrbach M, Venugopalan S, Guadarrama S, Saenko K, Darrell T. Long-Term Recurrent Convolutional Networks for Visual Recognition and Description // IEEE Trans Pattern Anal Mach Intell. – IEEE, 2017. – С. 130–141.
6. Tran Du, Bourdev L. Learning Spatiotemporal Features with 3D Convolutional Networks // 2015 IEEE International Conference on Computer Vision (ICCV). – IEEE, 2015. – С. 4489–4497.
7. Yue-Hei Ng J., Hausknecht M Beyond Short Snippets: Deep Networks for Video Classification // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2015. – С. 4694–4702.
8. Simonyan K, Zisserman A Very Deep Convolutional Networks for Large-Scale Image Recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – IEEE, 2014. – С. 3578–3584.

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССОВ В СИСТЕМАХ УПРАВЛЕНИЯ ПРОЕКТАМИ

Мельников Д. А., Петрова А. А.

Кафедра инструментального и прикладного программного обеспечения,
МИРЭА – Российский технологический университет

Институт информационных технологий

Москва, Российская Федерация

E-mail: dmbox2019@gmail.com, petrova@mirea.ru

В статье исследованы возможности использования нейронных сетей в системах управления проектами для оптимизации процессов принятия решений, повышения точности распределения ресурсов на примере прогнозирования матрицы ответственности RACI в зависимости от изменяющихся условий и требований к проекту. Рассмотрены примеры существующих готовых решений, использующих технологии искусственного интеллекта. Также описаны преимущества применения нейронных сетей в современных программных продуктах.

ВВЕДЕНИЕ

На сегодняшний день успешное управление проектами является достаточно сложной задачей для многих предприятий любого масштаба. Обработка больших объемов информации, анализ прошлых операций и адаптация под современные тенденции являются серьезными задачами для каждой компании, участвующей в проектной деятельности. Поэтому с каждым годом руководители проектов все чаще стараются внедрить в свою деятельность виртуальных помощников и другие программные продукты, основанные на нейронных сетях, для оптимизации своей деятельности.

I. ОСНОВНАЯ ЧАСТЬ

Одной из первоначальных задач руководителя проекта является распределение полномочий и ответственности между сотрудниками. Одним из самых актуальных подходов на сегодняшний день является матрица ответственности RACI, которая распределяет роли относительно решаемых задач. Несмотря на то, что такая матрица ответственности позволяет показать сбалансированную картину работы, тем самым повышая производительность труда, руководитель проекта может сталкиваться с проблемами перегруженности задач у одного или нескольких сотрудников или отсутствием ответственных участников за результат. Внесение корректировок может занять критическое для проекта количество времени и отрицательно отразиться на итоговых решениях. Если, к примеру, взять построение матрицы ответственности RACI для проекта, в котором задействовано несколько десятков участников, то необходимо подходить к распределению ролей с четким знанием всех профессиональных навыков каждого из участников и брать во внимание загруженность сотрудников в других проектах. Использование возможностей нейронной сети может упростить трудоемкую работу, исходя из анали-

за прошлого опыта участников, смоделировать все возможные варианты распределения ролей и найти баланс в делегировании задач. Обязанностями руководителя проекта останутся лишь проверка целесообразности предлагаемого решения, внесение необходимых корректировок в представленный нейросетью вариант и утверждение итогового результата.

Для определения точных зон ответственности каждого участника следует основываться на условных обозначениях матрицы RACI [1]:

- Исполнитель (R) – сотрудник, отвечающий за выполнение поставленной задачи;
- Ответственный (A) – сотрудник, принимающий или отвергающий предложения;
- Консультант (C) – сотрудник, консультирующий и согласовывающий принимаемые решения;
- Информированный (I) – сотрудник, уведомляемый о выполнении конкретной задачи.

Стоит учитывать, что вышеописанные зоны ответственности должны основываться на собранных ранее данных о профессиональных навыках сотрудников, включая компетенции, опыт работы, сертификация. Загруженность сотрудников можно оценивать, например, на основе количества задач на текущий момент или на основе обратной связи от руководителя. Вся эта информация предоставляется из базы знаний, которая хранит актуальные данные и регулярно обновляется компанией. Нейронная сеть может обучиться на этих данных и выполнять регрессии и классификации [2], чтобы прогнозировать матрицу RACI или определенную роль (R, A, C или I) для каждой задачи и сотрудника.

В качестве примера использования нейронной сети для распределения задач по матрице RACI воспользуемся одной из самых известных на сегодняшний день нейронных сетей ChatGPT версии GPT-4, разработанной компанией OpenAI [3].

Пример запроса для чат-бота представлен на рисунке 1, где описываются входные данные для нейронной сети, исходя из которых должна быть сформирована матрица RACI и распределены задачи по сотрудникам.

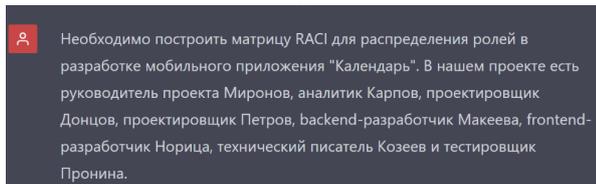


Рис. 1 – Запрос на построение матрицы RACI для разработки нового проекта

Исходя из запроса, проект был декомпозирован нейросетью на конкретные задачи, представленные на рисунке 2. Стоит заметить, что нейронная сеть самостоятельно определила этапы жизненного цикла для реализации данного проекта.

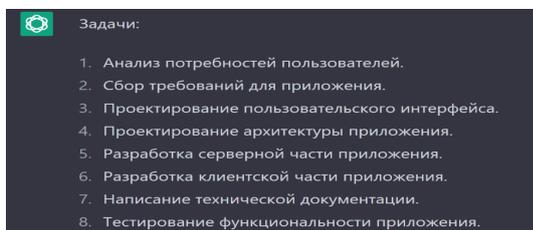


Рис. 2 – Декомпозиция задач нейросетью

Вместе с задачами нейросеть смогла построить матрицу RACI, продемонстрированную на рисунке 3. Стоит взять во внимание, что были учтены профессиональные навыки отдельных сотрудников.

Задача	Миронов	Карпов	Донцов	Петров	Макеева	Норича	Козеев	Пронина
1. Анализ потребностей пользователей	R	A	C	I	I	I	I	I
2. Сбор требований для приложения	R	A	C	I	I	I	I	I
3. Проектирование пользовательского интерфейса	I	R	A	A	I	I	I	I
4. Проектирование архитектуры приложения	I	I	R	A	I	I	I	I
5. Разработка серверной части приложения	I	I	I	R	A	I	I	I

Рис. 3 – Фрагмент построенной матрицы RACI

Исходя из того, что на данный момент отсутствует база знаний компании, то нейросеть не может определить профессиональные навыки отдельных сотрудников и распределяет задачи, исходя из их должностей. Интеграция такой нейронной сети с существующей системой управления проектами позволит автоматически обновлять матрицу RACI на основе изменений в проекте. Подобные решения уже используются в современных системах управления проектами:

- Smart Projects, которая обеспечивает поддержку полного цикла управления проектами и использует такие технологии, как

мультиагентные системы, онтологии и сетевые подходы для построения сложных систем планирования и управления;

- Auroga, ориентированная на создание оптимальных календарно-сетевых графиков проектов, использующая технологии искусственного интеллекта для применения правил и знаний планирования.

Применение современных программных продуктов, основанных на использовании искусственного интеллекта в проектах, позволяет:

- Вести подсчет вероятности возникновения рисков;
- Оказывать поддержку в сборе отчетности и построение план-графиков по проектам;
- Предоставлять информацию по проекту и промежуточным контрольным точкам;
- Устранять такие проблемы, как несоответствие, дублирование, двусмысленность;
- Распределять задачи по ролям, исходя из прошлого опыта участников проекта;
- Выбрать оптимальное решение, исходя из анализа базы знаний компании.

Однако стоит учитывать, что отдельное использование методологии RACI может не подойти для крупных проектов с большим количеством задач и участников, но использование этого подхода в совокупности с нейронной сетью может стать решением ряда проблем, возникающих при управлении проектами.

II. ЗАКЛЮЧЕНИЕ

Таким образом, при грамотном использовании возможностей нейронных сетей, безусловно, повышается эффективность и упрощается управление жизненным циклом в проектной деятельности, но и одновременно это не может заменить соответствующего специалиста, так как искусственный интеллект может не учитывать какие-то тонкости и особенности требований. Также не стоит забывать про человеческий фактор и динамически меняющиеся условия, которые по объективным причинам искусственный интеллект не может отслеживать.

СПИСОК ЛИТЕРАТУРЫ

1. Ганиходжаев Б. Ф., Аракелян Л. А. Формирование проектной команды и современные методы управления человеческими ресурсами / Б. Ф. Ганиходжаев, Л. А. Аракелян // Образование, наука и инновационные идеи в мире. - 2023. - №25(2). - С. 148-155.
2. Нейронные сети в машинном обучении [Электронный ресурс]. Режим доступа: <https://vr-app.ru/blog/neironnye-seti-v-masinnom-obucenii/>. Дата доступа: 20.10.2023.
3. Савранчук, Д. С. Искусственный интеллект по обработке естественного языка / Д. С. Савранчук // Молодой исследователь: от идеи к проекту : Материалы VII студенческой научно-практической конференции, Йошкар-Ола, 24–28 апреля 2023 года / Отв. редактор Д. А. Михеева. – Йошкар-Ола: Марийский государственный университет, 2023. – С. 130-132.

МОДЕЛИ И МЕТОДЫ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Михальков М. Д.

Кафедра программного обеспечения информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: solution.m@mail.ru

В публикации рассматриваются вопросы, связанные с построением интеллектуальных систем поддержки принятия решений и актуальностью предлагаемой задачи. Рассматриваются два основных класса методов построения интеллектуальных систем поддержки принятия решений и классы прикладных задач. Также рассматривается пример модели интеллектуальной системы поддержки принятия решений, которая объединяет преимущества эмпирических и аналитических методов с целью компенсации их недостатков.

ВВЕДЕНИЕ

Интеллектуальные системы поддержки принятия решений – любое ПО, которое оказывает поддержку пользователю в принятии решения, но не решает за человека.

Актуальность рассматриваемой темы можно описать следующим примером. Представим оператора некоего ситуационного центра. Этому оператору поступает звонок. Оператор регистрирует вызов, и его основная задача – понять из слов звонящего, какое конкретное происшествие произошло в реальном мире. Очевидно, что то, о чем говорит звонящий, и реальное происшествие – вещи связанные, но не эквивалентные. Соответственно, имеет место некая субъективность относительно данной ситуации, ведь звонящих может быть одновременно несколько: один увидел дым, другой – огонь, третий – бегущих людей, однако очевидно, что все они говорят об одном и том же. И вот задача этого оператора на основании полученных сообщений построить какую-то определенную картинку мира и запустить план реагирования, соответствующий данной ситуации. Очевидно, что поступающая информация может иметь совершенно разные источники: сообщения граждан, показатели датчиков, информация с различных систем мониторинга. Тут же встает задача охвата всех известных факторов рассматриваемой ситуации. Эта задача является довольно трудоемкой, поэтому в качестве решения можно рассматривать интеллектуальные системы поддержки принятия решений. Данные системы позволяют на основании всех имеющихся факторов сформировать общую ситуационную картину и предложить оператору наиболее рациональные варианты действия. Важно помнить, что данная система не принимает решений за человека, а только предоставляет возможные альтернативы.

В современной мире существует огромное множество сфер применения интеллектуальных систем поддержки принятия решений, таких как

общественная безопасность, промышленность, медицина, электронная коммерция и другие.

Во всех этих сферах прослеживаются следующие общие черты: крайне ограниченный промежуток времени на принятие решения (в основном на классификацию события), ограничения в принятии решений (регулируются нормативной базой), высокая цена за неверно принятое решение (как пример, жизни людей) [1].

I. КЛАССЫ МЕТОДОВ ПОСТРОЕНИЯ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Существует два основных класса методов построения интеллектуальных систем поддержки принятия решений: аналитические и эмпирические.

В случае использования аналитических методов разработчик должен построить определенную модель предметной области, в которой происходят инциденты, модель самих инцидентов и модель способов реагирования на них. В последующем он должен явно запрограммировать шаблоны поведения на конкретные ситуации в контексте множества входных данных и предметной области.

К примерам класса аналитических методов относятся императивное программирование, имитационное моделирование, деревья решений, автоматизация логического вывода и другие.

Положительным результатом использования методов этого класса решений является то, что результат может быть проверен (доказан) с использованием уже имеющегося математического аппарата, в одинаковых условиях гарантированы одинаковые решения, алгоритм можно нормативно утвердить.

Однако имеется и ряд отрицательных сторон, таких как: составлять алгоритмы и управлять ими долго и сложно, не всегда можно понять суть явления аналитически.

Вторым основным классом методов построения интеллектуальных систем поддержки при-

нятия решений является класс эмпирических методов. Общая черта этих методов состоит в том, что не моделируется суть происходящих событий, а строится модель, которая описывает множество входных наборов, которые показывают, что конкретно произошло, и множество выходных наборов, которые описывают реакции на входные наборы. Соответственно, при должной входной выборке система способна обучиться и выдавать выходные наборы с очень большой доверительной вероятностью.

Примерами класса эмпирических методов служат машинное обучение (в том числе нейросети, глубокое обучение, метод опорных векторов и другие), генетические алгоритмы.

Положительной стороной использования является тот факт, что достаточно качественной обучающей выборки или критерия правильности принятия решений.

К отрицательным результатам относятся: отсутствие гарантии правильности принятия решений, невозможность нормативного подкрепления.

II. КЛАССЫ ПРИКЛАДНЫХ ЗАДАЧ

Если организовать работу ситуационного центра согласно классу эмпирических методов, то в результате получается модель черного ящика, в котором в качестве входных воздействий выступает набор ситуаций, а в качестве выходных – полученный набор решений. В данном случае основная задача классификации – определить, что это за объект или явление, затем применить некое решение, заранее выработанное для такого класса явлений. Есть и более сложные ситуации, например – определение наличия причинно-следственной связи между двумя и более явлениями [2].

Соответственно, выделяют два класса прикладных задач: задачи, толерантные к ошибке, и задачи, не толерантные к ошибке.

В случае задач, толерантных к ошибке, имеет место задачи формирования персональных рекомендаций, задачи коммуникации и связи, задачи, связанные с электронной коммерцией. В этом случае допустим некоторый процент ошибочных решений при условии, что общий результат положителен. Кроме того, в данной группе задач нет однозначного нормативного регулирования принимаемых решений.

В случае задач, не толерантных к ошибке, имеет место задачи, связанные с медициной, промышленностью, обеспечением безопасности. В данном случае недопустим даже единичный случай ошибочного решения, решения должны приниматься с учетом нормативной базы, присутствует личная ответственность оператора за ошибку.

III. ПРИМЕР МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Анализируя описанные классы методов построения интеллектуальных систем поддержки принятия решения и классы прикладных задач, получаем с одной стороны, что нужна возможность проверить и обосновать любое решение, необходимы рациональные основания для принятия каждого решения, недопустимы ошибки; а с другой стороны, нет возможности аналитически описать всю логику принятия решений во всех возможных классах ситуаций.

Решением может служить объединение преимуществ эмпирических и аналитических методов для того, чтобы скомпенсировать их недостатки. Тем самым может быть получена гибридная модель интеллектуальной системы поддержки принятия решений [3].

Рассмотрим общий алгоритм функционирования такой системы.

В начале происходит обработка событий оператором, формируется история, состоящая из блоков ситуация – решение. Затем алгоритм формирует первичные гипотезы о том, почему оператор принял то или иное решение в каждой конкретной ситуации. В процессе анализа историй событий гипотезы подтверждаются или опровергаются. Далее алгоритм обобщает подтвержденные гипотезы. В результате получается читаемый для человека набор гипотез, которые эксперт просматривает и утверждает или опровергает их [4].

ЗАКЛЮЧЕНИЕ

Объединение преимуществ эмпирических и аналитических методов для того, чтобы скомпенсировать их недостатки, позволяет получить модель белого ящика, в котором в качестве входных наборов выступает множество ситуаций, а в качестве выходных наборов – множество решений и обоснований. Это позволяет гарантировать принятие одинаковых решений в одинаковых ситуациях, такое решение обосновано и подтверждено, всегда можно установить ответственного за любой шаг в принятии решения.

1. Аксенов, К. А. Моделирование и принятие решений в организационно-технических системах: учебное пособие. В 2 ч. Ч.1 / К. А. Аксенов, Н. В. Гончарова. – Екатеринбург: Изд-во Урал. ун-та, 2015. – 104 с.
2. Заболотский, В. П. Математические модели в управлении: учебное пособие / В. П. Заболотский, А. А. Оводенко, А. Г. Степанов — СПбГУАП. СПб., 2001. – 196 с.: ил.
3. Кизим, Н. А. Адаптивные модели в системах принятия решений: Монография / Под ред. Н. А. Кизима, Т. С. Клебановой. – Х.: ИД «ИНЖЭК», 2007. – 368 с. Русск. яз.
4. Никонов, О. И. Математическое моделирование и методы принятия решений: учеб. пособие / О. И. Никонов, С. В. Кругликов, М. А. Медведева. – Екатеринбург: Изд-во Урал. ун-та, 2015. – 100 с.

МОДУЛЬ ОБРАБОТКИ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ СООБЩЕНИЙ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Никифоров С. А., Садовский М. Е., Гойло А. А.
Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Кафедра теории и практики перевода, Минский государственный лингвистический университет
Минск, Республика Беларусь
E-mail: nikiforov.sergei.al@gmail.com, sadovski@bsuir.by, artemgoylo@gmail.com

В работе предлагается подход к реализации модуля обработки естественно-языковых сообщений систем, построенных по Технологии OSTIS (ostis-систем).

ВВЕДЕНИЕ

Технология OSTIS использует набор внешних языков представления знаний, а именно SCg-код, SCs-код и SCn-код.[1]

При этом остается нерешенной проблема взаимодействия ostis-систем с неподготовленным пользователем, не обладающим навыками работы с данными внешними языками. По этой причине является актуальной реализация возможности взаимодействия с системой на естественном языке.

Таким образом, целью данной работы является разработка подхода к реализации начальной версии естественно-языковых интерфейсов ostis-систем.

I. ПРЕДЛАГАЕМЫЙ ПОДХОД

Разработка модуля предполагает работу по двум направлениям:

- Трансляция структур базы знаний в естественно-языковое представление;
- Обработка естественно-языковых запросов к системе.

Трансляция структур базы знаний в естественно-языковое представление

В рамках данного направления реализуется подмодуль, целью которого является трансляция конструкций в базе знаний в тексты естественного языка, которые могут содержать вставки SCn-кода. Результатом работы данного подмодуля является файл ostis-системы с размеченным посредством Latex текстом, который содержит макросы для описания SCn-конструкций.[2]

В настоящее время видится 2 основных варианта использования данного подмодуля:

- Выгрузка произвольного фрагмента БЗ в естественно-языковое представление;
- Навигация по базе знаний на естественном языке в SC-web.[3]

В первом случае выбор структуры ложится на пользователя: он может сформировать новую, содержащую необходимые ему фрагменты БЗ, или выбрать существующую, например, – некоторую предметную область.

Соответственно, в таком случае предлагаемый вариант работы данного подмодуля включает в себя следующие этапы:

- Разбиение транслируемой структуры на ориентированное множество подструктур;
- Трансляция каждой из полученных подструктур в естественно-языковое представление.

Во втором случае (в режиме навигации) структуру требуется формировать автоматически. При навигации по базе знаний в SCn или SCg на экран выводится семантическая окрестность элемента. Соответственно, перед приведенными выше двумя этапами добавляется третий – формирование структуры, содержащей семантическую окрестность.

Таким образом, для работы подмодуля требуются следующие агенты:

- Агент поиска семантической окрестности;
- Агент разбиения структуры на ориентированное множество подструктур;
- Агент трансляции структуры в естественно-языковое представление.

Агент поиска семантической окрестности формирует структуру из элементов, составляющих семантическую окрестность некоторого элемента.

Агент разбиения структуры на ориентированное множество подструктур разбивает транслируемую структуру на ориентированное множество структур, являющееся ее разбиением.

Порядок структур во множестве задает последовательность структур при трансляции.

Построение данного множества подструктур может быть выполнено как по предварительно заданной спецификации структуры, так и алгоритмически.

При этом для обеспечения возможности трансляции произвольной структуры остается необходимость в алгоритмическом варианте. Для определения порядка подструктур могут быть использованы:

- Ролевые отношения от структуры к ее элементам;
- Зависимости между элементами структуры.

В первом случае между различными отношениями в БЗ требуется задать порядок, на основе которого осуществляется выбор семантических окрестностей элементов, входящих в более приоритетные структуры.

Во втором случае предлагается выполнение поиска зависимостей между понятиями исходя из отношений, которыми они связаны, и последующее построение соответствующего порядка подструктур.

Агент трансляции структуры в естественно-языковое представление в свою очередь включает в себя:

- Агент трансляции структур в базе в промежуточную версию естественно-языкового представления;
- Агент преобразования промежуточной версии естественно-языкового представления в финальное.

Входом **агента трансляции структур в базе в промежуточную версию естественно-языкового представления** является транслируемая структура, выходом – файл *ostis*-системы с полученным промежуточным текстом, который может включать нарушения нормы естественного языка (например, неправильные окончания).

Агент преобразования промежуточной версии естественно-языкового представления в финальное предполагается реализовать с применением языковой модели. Входным и выходным аргументами данного агента являются файлы *ostis*-системы. Входной файл содержит преобразуемый текст, выходной – результат преобразования.

Естественно-языковые запросы к системе

В рамках данного направления реализуется подмодуль, целью которого является обработка естественно-языковых запросов к системе из типового набора с возможностью его расширения.

В его состав входят:

- Агент классификации сообщений;
- Агент ответа на сообщения;
- Агент фильтрации структуры.

Агент классификации сообщений предполагается создать на основе одного из существующих классификаторов – Rasa[4], Wit AI[5] и т. д.

Агент ответа на сообщения принимает на вход классифицированное сообщение, а его выходом является файл *ostis*-системы с текстом в формате, описанном в начале предыдущего раздела. Работа агента сводится к двум основным этапам:

- Попытка ответа по правилам;

- Поиск соответствующего вопросу действия для ответа (если первый этап завершен не успешно).

На первом этапе для ответа по правилам предлагается использовать агент ответа на сообщение из НИКА.[6] Цель введения этих правил – получение ответов с наиболее близкими к естественному языку формулировками на часто задаваемые вопросы (например, «Кто ты?» и т. п.).

Если по правилам ответ не получен, выполняется поиск классов действий, связанных с классом сообщения отношением *ответное действие**, создается и иницируется экземпляр данного действия с аргументом из сообщения.

Полученная структура (ответ действия) может быть протранслирована в естественно-языковое представление с использованием подмодуля, описанного в предыдущем разделе. Результат такой трансляции и будет являться ответом на сообщение.

Для того, чтобы исключить возможность генерации слишком громоздкой структуры вводятся **агент фильтрации структуры**, который обрабатывает структуру до ее трансляции в естественно-языковое представление. Агент принимает на вход фильтруемую структуру, а его выходом является некоторое ее подмножество.

Фильтрацию предлагается осуществлять на основе шаблонов, множество которых соответствует некоторому классу действий.

II. ЗАКЛЮЧЕНИЕ

Был разработан подход к реализации начальной версии модуля обработки естественно-языковых сообщений *ostis*-систем, предоставляющий возможность:

- Транслировать структуры базы знаний в естественно-языковое представление;
- Обрабатывать естественно-языковые запросы к системе.

1. Технология комплексной поддержки жизненного цикла семантически совместимых интеллектуальных компьютерных систем нового поколения / под общ. ред. В. В. Голенкова. – Минск : Бестпринт, 2023. – 1064 с.
2. SCn Latex plugin [Электронный ресурс] – Режим доступа: <https://github.com/ostis-ai/scn-latex-plugin>. – Дата доступа: 15.10.2023.
3. SC-web [Электронный ресурс] – Режим доступа: <https://github.com/ostis-ai/sc-web>. – Дата доступа: 15.10.2023.
4. Rasa Platform [Электронный ресурс] – Режим доступа: <https://rasa.com/>. – Дата доступа: 17.10.2023.
5. Wit AI [Электронный ресурс] – Режим доступа: <https://wit.ai/>. – Дата доступа: 17.10.2023.
6. НИКА [Электронный ресурс] – Режим доступа: <https://github.com/ostis-apps/nika>. – Дата доступа: 18.10.2023.

МОДЕЛИ ГРАФОВЫХ БАЗ ДАННЫХ ДЛЯ АНАЛИЗА ПРЕДПОЧТЕНИЙ ПОЛЬЗОВАТЕЛЕЙ

Павлюченко К. А., Киселев А. И.

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

E-mail: mendorill@mail.ru, and.kis135@gmail.com

Модели графовых баз данных предоставляют мощный инструмент для анализа предпочтений пользователей. Их использование позволяет эффективно хранить и анализировать связи между объектами информационных систем, что может быть весьма полезным бизнесу для повышения заинтересованности пользователей и увеличения доходности. Данный исследовательский проект изучает, как графовые базы данных могут быть применены для анализа предпочтений пользователей и создания персонализированных рекомендаций.

ВВЕДЕНИЕ

Наличие возможности детального анализа предпочтений пользователей имеет существенное влияние на доходность предприятий в современном мире. Знание о том, что предпочитают пользователи, позволяет создавать персонализированные рекомендации, улучшать поисковую выдачу и повышать качество обслуживания. В данной работе рассматривается применение моделей графовых баз данных для анализа предпочтений пользователей, исследуется их реализация и полученные результаты.

Модели графовых баз данных предоставляют удобный и эффективный способ хранения и анализа данных, организованных в виде графа [1]. Модели графовых баз данных имеют следующие преимущества:

- Гибкость и масштабируемость создаваемых моделей;
- Интуитивно понятное моделирование сложных связей;
- Широкие возможности для глубокого анализа данных.

Граф представляет собой совокупность узлов и связей между ними. В контексте анализа предпочтений пользователей информационной системы, узлами могут быть аккаунты пользователей и другие объекты системы, а связи между ними будут отражать предпочтения конкретных пользователей [2].

I. МОДЕЛИРОВАНИЕ ГРАФОВОЙ БАЗЫ ДАННЫХ

Моделирование графовой базы данных начинается с анализа и получения доступа к источнику данных. Это включает в себя изучение документации и получение ключей публичного API сервиса. Если сервис не предоставляет пользовательских API, то для получения данных можно применить техники веб-скрейпинга или же использовать сторонние API для доступа к контенту веб-ресурса. Исходя из анализа структуры доступных данных происходит создание графовой модели, определяются цели и параметры сбо-

ра данных, (например, информация о подписках, лайках, комментариях, просмотрах и других активностях пользователей). В полученной модели каждый пользователь и объект (например, фотография или видео) представлены узлами, а связи между ними отражают действия пользователя по отношению к объектам (например, "пользователь-лайк-фото"). С использованием полученного доступа начинается сбор данных о пользователях, их профилях, подписках, а также данных о контенте, который представляет наибольший интерес (например, фотографии, видео, текстовые публикации). Используя полученную графовую базу данных, становится возможно определить, какие объекты наиболее популярны среди пользователей, какие пользователи имеют схожие интересы и какие темы наиболее обсуждаются. Этот анализ позволяет лучше понимать предпочтения пользователей.

II. РАЗРАБОТКА РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ

На основе моделей анализа данных разрабатываются рекомендательные системы [3], которые предоставляют пользователям персонализированный контент и рекомендации на основе их предпочтений. Это может включать в себя рекомендации для просмотра видео, фотографий, текстовых публикаций и других объектов системы.

Для удобства описания данных, используется язык запросов, который поддерживает графовую базу данных Neo4j [4]. Пример запроса для создания узлов пользователей представлен на рисунке 1.

```
graph.run("CREATE (u1:User {name: 'User1'})")  
graph.run("CREATE (u2:User {name: 'User2'})")
```

Рис. 1 – создание узлов пользователей

Пример создания узлов категорий представлен на рисунке 2.

```
graph.run("CREATE (c1:Category {name: 'Category1'})")
graph.run("CREATE (c2:Category {name: 'Category2'})")
```

Рис. 2 – создание узлов категорий

Создание связей между пользователями и их предпочитаемыми категориями продемонстрировано на рисунке 3.

```
graph.run("MATCH (u:User), (c:Category)
WHERE u.name = 'User1' AND c.name = 'Category1'
CREATE (u)-[:LIKES]->(c)")

graph.run("MATCH (u:User), (c:Category)
WHERE u.name = 'User1' AND c.name = 'Category2'
CREATE (u)-[:LIKES]->(c)")

graph.run("MATCH (u:User), (c:Category)
WHERE u.name = 'User2' AND c.name = 'Category1'
CREATE (u)-[:LIKES]->(c)")
```

Рис. 3 – создание связей между пользователями и их предпочитаемыми категориями

Полученная модель может быть представлена в виде графа, изображённого на рисунке 4.

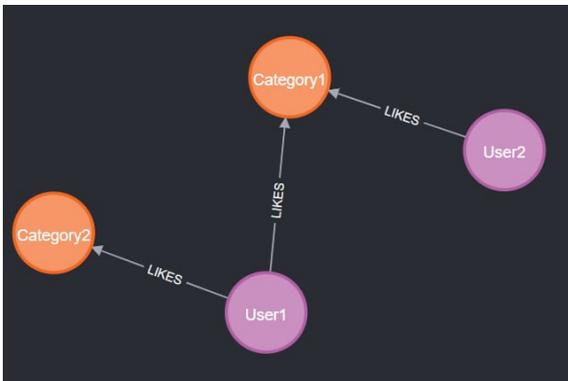


Рис. 4 – граф базы данных

Продемонстрировать работу созданной графовой модели можно на примере выполнения запросов для анализа предпочтений пользователей. Один из часто используемых запросов - поиск пользователей с общими предпочтениями. Можно найти всех пользователей, которые имеют общие предпочтения с «User1», запрос представлен на рисунке 5.

```
result = graph.run(
  "MATCH (u1:User {name: 'User1'})-[:LIKES]->(c:Category) "
  "MATCH (u2:User)-[:LIKES]->(c) "
  "WHERE u2 <> u1 " "RETURN u2.name, "
  "COLLECT(DISTINCT c.name) AS common_likes")

for record in result:
  print(f"{record['u2.name']} shares common likes with "
  "User1: {'', '.join(record['common_likes'])}")
```

Рис. 5 – нахождение всех пользователей, имеющих общие предпочтения с «User1»

Результатом выполнения данного запроса будет «User2 shares common likes with User1: Category1». Другие типы исследуемых запросов включают в себя анализ сетей, выявление лидеров мнений или создание рекомендаций на основе предпочтений.

III. Выводы

При реализации анализа предпочтений пользователей важно учесть оптимизацию запросов. Эффективные запросы позволяют быстро анализировать большие объемы данных. Для этого можно использовать кэширование и оптимизацию запросов к базе данных.

Помимо базовых запросов, можно рассмотреть применение более сложных аналитических методов, таких как выявление общих паттернов и предсказательное моделирование. Эти методы могут дополнительно улучшить результаты и эффективность анализа предпочтений пользователей.

Реализация анализа предпочтений пользователей с использованием графовых баз данных позволяет наглядно и эффективно моделировать связи между пользователями и объектами анализируемой системы. После успешной реализации и оптимизации анализа предпочтений пользователей, этот подход может быть интегрирован в реальные проекты и приложения в качестве рекомендательной системы для социальных сетей, интернет-магазинов и других сервисов. Полученные в ходе работы над исследовательским проектом результаты могут быть использованы для значительного повышения заинтересованности пользователей и увеличения доходности бизнеса, путём создания персонализированных рекомендаций.

СПИСОК ЛИТЕРАТУРЫ

1. Применение графовых баз данных [Электронный ресурс] / – Режим доступа: <https://nitrosdata.ru/2019/02/20/primenenie-grafovyh-baz-dannyh>. – Дата доступа: 11.10.2023.
2. Графовый анализ — обзор и области применения [Электронный ресурс] / – Режим доступа: <https://habr.com/ru/companies/glowbyte/articles/594221>. – Дата доступа: 13.10.2023.
3. Hybrid Recommender Systems: Survey and Experiments / Robin Burke. – User Modeling and User-Adapted Interaction 12(4), 2002.
4. Neo4j DBMS Documentation [Электронный ресурс] / – Режим доступа: <https://neo4j.com/docs/>. – Дата доступа: 12.10.2023.

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПСИХОТЕРАПИИ

Радченко А. Д.

Кафедра вычислительных методов и программирования,
Белорусский государственный университет информатики и радиоэлектроники
Объединённый институт проблем информатики Национальной академии наук Беларуси
Минск, Республика Беларусь
E-mail: a.radchenko@bsuir.by

В данной статье обзревается примеры использования технологий с искусственного интеллекта в области охраны психического здоровья

ВВЕДЕНИЕ

Искусственный интеллект и нейросети – это технологии, которые с каждым годом становятся все более распространенными и востребованными. Первое, что нужно отметить – нейросети и искусственный интеллект могут использоваться для анализа и обработки большого объема данных, связанных с психологическим здоровьем. Одним из примеров использования нейросетей в психологии – исследование, проведенное в Университете Цюриха. В этом исследовании использовались нейросети для анализа данных, связанных с депрессией. Было обнаружено, что нейросети могут помочь предсказать вероятность того, что у человека разовьется депрессия в будущем. Еще одним примером использования искусственного интеллекта в психологии является использование роботов в терапии. Некоторые исследования показывают, что роботы могут быть эффективным средством для терапии детей с аутизмом. Роботы могут использоваться для обучения детей различным социальным навыкам, таким как установление контакта с другими людьми, распознавание эмоций и развитие коммуникативных навыков. Нейросети и искусственный интеллект могут использоваться для создания приложений, которые помогают людям справляться с психологическими проблемами. Еще одной областью, в которой нейросети и искусственный интеллект могут помочь в психологии, является диагностика и лечение психических расстройств. Например, нейросети могут использоваться для анализа изображений мозга и помочь в диагностике расстройств, таких как шизофрения, депрессия и биполярное расстройство.

I. ОБЗОР СУЩЕСТВУЮЩИХ ПРОДУКТОВ С ПРИМЕНЕНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПСИХОЛОГИИ И ПСИХОТЕРАПИИ

Уже сейчас экспертные системы, приложения, чат-боты и компьютерные игры в сфере психологической помощи успешно используют искусственный интеллект. Так например Watson Health – это когнитивная система с поддержкой искусственного интеллекта, выполняет функции медицинского консультанта и эксперта. Она способна

не только генерировать колоссальные объемы медицинской информации, но и распознавать малейшие отклонения в состоянии здоровья пациентов, которые может упустить даже высококвалифицированный врач с многолетним опытом[1]. Чтобы помочь врачам в лечении своих пациентов, после того как врач задал системе запрос с описанием симптомов и других сопутствующих факторов, Watson сначала анализирует входные данные, чтобы определить наиболее важные фрагменты информации; затем собирает данные о пациенте, чтобы найти факты, относящиеся к медицинской истории пациента; затем изучает доступные источники данных для формирования и проверки гипотез и, наконец, предоставляет список индивидуальных рекомендаций, оцениваемых с точки зрения достоверности. Источники данных, которые Watson использует для анализа, могут включать рекомендации по лечению, данные электронной медицинской карты, заметки медицинских работников, исследовательские материалы, клинические исследования, журнальные статьи и информацию о пациентах.

Робот RP-VITA используется для обеспечения удаленной связи между медицинскими работниками и пациентами. Он дистанционно контролирует самочувствие пациентов, получая доступ к их медицинским записям. Данная технология обеспечивает поддержку при проведении психологических и неврологических оценок, а также обследований в условиях интенсивной терапии[2].

Помимо этого существует экспертная система способная ставить диагнозы и предлагать методы лечения в психиатрии. Для этого она использует способности искусственного интеллекта кодировать экспертные знания о психических расстройствах. Такие системы способны предложить методы лечения учитывая потребности пациента, согласования планов лечения, бюджета и других проблем со здоровьем.

Существует также ряд приложений и чат-ботов направленных на эмоциональную поддержку, в их основе лежит чат-бот на основе текстовой нейросети GPT. Она обучается на сообщениях пользователя и пытается адаптироваться под собеседника. Ниже приведены примеры таких ботов:

- Youper – персонализированный помощник, способный уменьшить симптомы стресса и депрессии.
- Replika – это чат-бот обеспечивающий эмоциональную поддержку при депрессивных или тревожных расстройствах.
- Tess – чат-бот предлагающий стратегии преодоления имеющихся психологических проблем, которые разрабатывает с использованием искусственного интеллекта.

В современной практике используются программы, направленные на решение различных психологических проблем. Так например в работе Ф. Фицпатрик и др. описан полностью автоматизированный разговорный агент для реализации программы самопомощи студентам колледжей, имеющим симптомы тревоги и депрессии [3]. А в работе Л. Афонсо с соавторами обсуждается разработка веб-приложения для онлайн-терапии, которое представляет собой интерактивную платформу, созданную на основе социальных сетей и направленную на восстановление психического здоровья [4].

В настоящее время в области охраны психического здоровья и лечебно оздоровительной терапии также используют компьютерные игры. Например терапевтические компьютерные игры помогают подросткам повысить уверенность в себе и совершенствовать навыки решения проблем [5]. Преимущество использования компьютерных игр заключается в повышенной вовлеченности пациентов. Технологии искусственного интеллекта можно помогут сделать игры, которые можно адаптировать к потребностям пациента и они могут использоваться для управления игровым процессом, чтобы пациент практиковал навыки в необходимых областях [6].

Кроме этого искусственный интеллект может имитировать практикующего врача с возможностями, превосходящими возможности человека. На сегодняшний день существуют передовые технологии, позволяющие оценить параметры пациентов, выходящие за пределы человеческих чувств, такие как:

- тепловидение для распознавания изменений температуры человека;
- оптическое считывание выражения лица и моргания глаз;
- анализ голосовых данных для восприятия тонких различий в речевых характеристиках;

Оценка физиологических особенностей например учащенное сердцебиение или изменение температуры при ответе на заданные вопросы во время терапии могут предоставить терапевту ценные дополнительные данные.

II. ЗАКЛЮЧЕНИЕ

Уже сейчас существует ряд успешных продуктов с использованием искусственного интеллекта применяющихся для помощи пациентам в области психологии, а также расширяющие возможности врачей. Данное направление разработок является перспективным не только с коммерческой точки зрения но и повышения уровня медицинского обслуживания. Одним из возможных направлений развития искусственного интеллекта в психотерапии может стать объединения сбора физиологических данных пациента и персональных данных о имеющихся проблемах со здоровьем с чат-ботами. Это позволит искусственному интеллекту быстрее реагировать на изменение состояния пациента и пытаться начать диалог с пользователем до того как человек сам заметит изменения состояния. Для этого необходимо создать систему автоматического сбора физиологических данных и их обмена и обработки искусственным интеллектом.

СПИСОК ЛИТЕРАТУРЫ

1. Когнитивная система IBM Watson Health – прорыв в сфере здравоохранения [Электронный ресурс] / Mentamore – Режим доступа: <https://mentamore.com/covremennyye-tekhnologii/kognitivnaya-sistema-ibm-watson-health-proryv-v-sfere-zdravooxraneniya.html>
2. Медицинский робот RP-VITA [Электронный ресурс] / Роботы и новые технологии – Режим доступа: <https://robroy.ru/medicinskij-robot-rp-vita.html>
3. Fitzpatrick F., Doherty A., Lacey G. Using artificial intelligence in infection prevention // Current Treatment Options in Infectious Diseases. –2020. –Vol. 12 (2). –P. 135–144.
4. Afonso L. C. S. A recurrence plot-based approach for Parkinson's disease identification / L. C. S. Afonso et al. // Future Generation Computer Systems. – 2019. – Vol. 94. – P. 282–292.
5. Coyle D. Personal investigator: a therapeutic 3D game for adolescent psychotherapy / D. Coyle et al. // Interactive Technology and Smart Education. – 2005. – Vol. 2 (2). – P. 73–88.
6. Fujita H., Wu I. C. A special issue on artificial intelligence in computer games: AICG // Knowledge-Based Systems. – 2012. – Vol. 34. – P. 1–2.

МОДЕЛИ И СРЕДСТВА ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ПЕРСОНАЛЬНОГО МЕДИЦИНСКОГО ОБСЛУЖИВАНИЯ

Сальников Д. А., Загорский А. Г., Крищенко В. А., Кондратьева И. Д., Корсакова С. К.
Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ,salnikov,a.zagorskij, krish@bsuir.by, kondrashka04, swe123ta@gmail.com

В работе описаны принципы и методики проектирования интеллектуальных систем персонального медицинского обслуживания, затрагивая аспекты приватности, этики, совместимости с техническими системами. Также сформированы аспекты, на которые должны опираться разработчики и системные инженеры при создании подобных систем.

ВВЕДЕНИЕ

Персональные ассистенты имеют потенциал стать центральной точкой взаимодействия пользователя с информацией и сервисами. Они могут стать универсальным интерфейсом, позволяющим интегрировать и использовать различные ресурсы и сервисы на унифицированной основе, учитывая индивидуальные характеристики пользователей. Разновидностью таких виртуальных ассистентов являются персональные медицинские ассистенты (ПМА) – программные системы или устройства, которые предназначены для персонализированной помощи людям в управлении и улучшении их здоровья и медицинских данных [1].

I. ПРОБЛЕМАТИКА СИСТЕМ ПЕРСОНАЛЬНОГО МЕДИЦИНСКОГО ОБСЛУЖИВАНИЯ

Некоторые медицинские ассистенты ограничены в работе только с определенными системами или устройствами, что затрудняет обмен информацией между разными поставщиками здравоохранения.

Использование ПМА не должно заменять врачей и медицинский персонал, а должно служить дополнением. Недостаток сотрудничества между разными специалистами и дисциплинами может создать проблемы в долгосрочном уходе и лечении пациентов.

Пациенты и медицинский персонал могут не иметь достаточного уровня знаний, чтобы правильно использовать медицинский ассистент и понимать предоставляемую им информацию. Это может привести к неправильному использованию или недопониманию результатов.

ПМА могут быть неполными или недостаточно точными при предоставлении медицинских советов или информации. Неверные рекомендации могут повлечь за собой негативные последствия для здоровья.

Для использования системы требуется доступ к интернету и соответствующей технологической инфраструктуре. Это может ограничить доступность ПМА для некоторых групп пациен-

тов, особенно для тех, кто живет в отдаленных или недостаточно технологически развитых регионах.

В использовании персонального медицинского ассистента возникают этические вопросы, связанные с сбором и использованием медицинских данных, а также с автономией пациента и конфиденциальностью.

II. ТРЕБОВАНИЯ К ПЕРСОНАЛЬНОМУ МЕДИЦИНСКОМУ АССИСТЕНТУ

Персональный медицинский ассистент должен обладать рядом ключевых характеристик и функций, чтобы эффективно служить пациентам и помогать им в управлении их здоровьем. Ниже приведены некоторые аспекты, которые определяют уровень качества персонального медицинского ассистента [2].

Интероперабельность и совместимость. Система должна иметь способность наладить взаимодействие с разными медицинскими системами и устройствами для обеспечения непрерывности поддержания здоровья.

Обучаемость. Использование технологий искусственного интеллекта позволяет ПМА расширять собственную базу знаний и повышать количество и качество сформированных навыков: анализ состояния пользователя, напоминания о приеме лекарств, визитах к врачу, механизмы связи с медицинскими специалистами, и так далее.

Персонализация. ПМА должен адаптироваться к уникальным потребностям каждого пациента, предоставляя персонализированные рекомендации и информацию.

Сбор и мониторинг медицинских данных. Система должна иметь возможность собирать и мониторить медицинские данные пациента, такие как вес, давление, уровень сахара в крови и другие, и предоставлять анализ и отчеты.

Предоставление полной информации пользователю. Ассистент должен предоставлять информацию о рисках, заболеваниях, лечении, лекарствах и здоровом образе жизни, анализировать большие объемы данных и выявлять тенденции и аномалии, а также иметь способность дообучать

пользователя на основе сформированной базы знаний.

III. ПРЕДЛАГАЕМЫЙ ПОДХОД

В качестве ИС рассмотрим системы на основе Технологии OSTIS [3]. Предлагаемый подход предполагает хранение в базах знаний систем онтологий и предметных областей, описывающих:

- модель пользователя;
- модель болезни;
- модель строения тела человека;
- модель лекарственных средств.

Реализация в базе знаний конкретных элементов формируется на основе онтологий. Это позволяет поддерживать семантическую совместимость базы знаний системы на высоком уровне, а также позволяет системам обмениваться информацией друг с другом.

Для хранения персональных данных используется база знаний персонального ostis-ассистента. Такая система основана на онтологии пользователя и является системой-интерфейсом для пользователя с остальными системами. Персональный ostis-ассистент способен взаимодействовать с другими системами и использовать их функционал для своих целей: анализ текущих данных и вывод результата пользователю, агрегирование для формирования статистики, формирование новых знаний на основе имеющихся.

В такую систему можно интегрировать различные методы диагностики. Пример структуры системы, использующей методы ФСД-диагностики, представлен на рисунке 1.

Основной целью создания такой ИС является выявление рисков развития заболеваний и латентных(скрытых) стадий их развития, а также для своевременной диагностики манифестных заболеваний, используя алгоритм:

- запись волнового сигнала организма и вычисление Добеши-3 вейвлетов;
- формирование вейвлет образа пациента и логико-статистическое распознавание диагностических маркеров;
- выявление информативных маркеров и логико-статистическое распознавание рисков патогенеза;
- формирование профиля рисков и логико-семантическое распознавание системных процессов;
- формирование профиля системных процессов и логико-семантическая генерация индивидуального сценария профилактического оздоровления.

IV. ЗАКЛЮЧЕНИЕ

Интеллектуальная система медицинского обслуживания предназначена для мониторинга здоровья с учетом интеграции информационных и программных систем. Использование Технологии OSTIS открывает наиболее прозрачные перспективы онтологического проектирования, производства и эксплуатации семантически совместимых гибридных интеллектуальных компьютерных систем мониторинга здоровья и других интеллектуальных медицинских систем.

СПИСОК ЛИТЕРАТУРЫ

1. Уотермен, Д. Руководство по экспертным системам: Пер. с англ. - М.: Мир, 1989.-388 с., ил.
2. Ростовцев, В. Н. Основы здоровья / В. Н. Ростовцев. – Мн.: Минсктиппроект, 2002. – 110 с.
3. Голенков, В. В. Открытая технология онтологического проектирования, производства и эксплуатации семантически совместимых гибридных интеллектуальных компьютерных систем / В. В. Голенков, Н. А. Гулякина, Д. В. Шункевич. – Минск : Бестпринт, 2021. – 690 с.

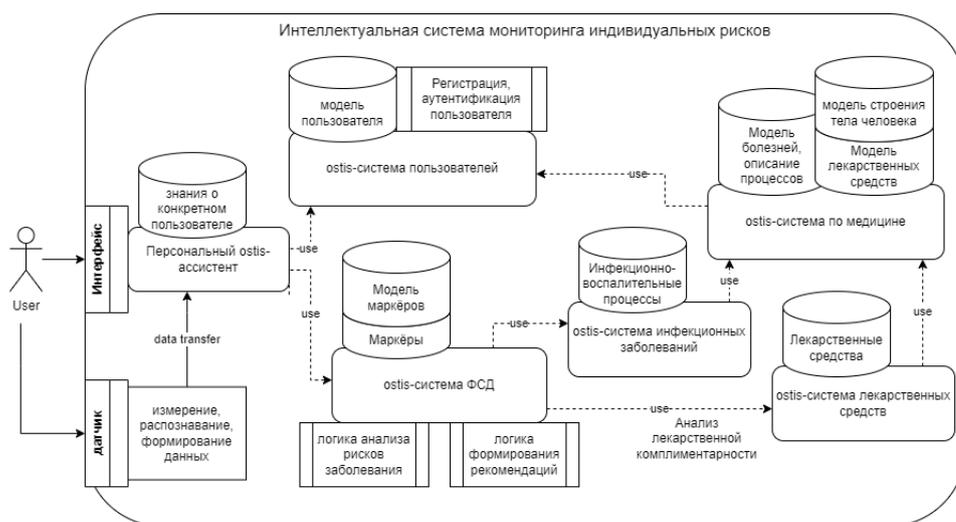


Рис. 1 – Структура интеллектуальной системы персонального медицинского обслуживания в виде взаимодействующих ostis-систем

МОБИЛЬНОЕ УСТРОЙСТВО ДЛЯ ДИАГНОСТИКИ ДИСФУНКЦИ ОПОРНО-ДВИГАТЕЛЬНОГО АППАРАТА ЧЕЛОВЕКА

Смык М. Н., Живицкий И. А.

Факультет компьютерного проектирования, Кафедра электронной техники и технологий,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: smykmiha@gmail.com

В тезисах рассмотрено спроектированное мобильное устройство для диагностики дисфункций опорно-двигательного аппарата человека и его актуальность. Описаны нарушения, которые устройство поможет диагностировать. Описаны задачи и цели устройства и его функционал. Рассмотрен принцип работы, а также отличительные черты данного устройства от аналогов. Описана программная составляющая устройства.

ВВЕДЕНИЕ

На сегодняшний день примерно 1,71 млрд. человек в мире страдают от нарушений и болезней костно-мышечной системы. Данные нарушения и болезни значительно ограничивают подвижность и моторику человека, приводя к преждевременному прекращению его трудовой деятельности, а также к сокращению возможностей для участия в жизни общества. Нарушения и болезни опорно-двигательного аппарата варьируются в широком диапазоне и занимают ведущее место среди факторов инвалидности в мире (17 процентов от общего количества инвалидностей, вызванных теми или иными заболеваниями).[1] Диагностика данных заболеваний является одной из наиболее приоритетных задач, так как своевременная диагностика, терапия и лечение позволяют замедлить и даже устранить нарушения на ранних этапах. Целями разрабатываемого устройства являются:

1. упрощение диагностики нарушений опорно-двигательного аппарата;
2. увеличение доступности проведения диагностики как в медицинских учреждениях, так и в домашних условиях;
3. расширение функционала уже существующих устройств;
4. увеличение количества исследуемых параметров для детальной диагностики.

I. ПРИНЦИП РАБОТЫ УСТРОЙСТВА

Прибор построен по принципу стабилметрической платформы, которая служит для анализа способности человека управлять позой тела и обеспечения биологической обратной связи по опорной реакции. Принцип действия разрабатываемого устройства основан на измерении вертикальных сил, прилагаемых к силоизмерительным датчикам и возникающих в результате размещения исследуемого объекта на опорной поверхности платформы, вычисления массы объекта и координат точки приложения равнодействующей силы, воздействующей со стороны объекта на

опорную поверхность платформы общего центра давления. Стабилометрия – это широкий спектр методических приемов, заключающихся в измерении координат центра давления, создаваемого человеком на плоскость опоры, в определенных условиях за определенный период времени, с целью количественной оценки двигательных возможностей или с целью создания биологической обратной связи по опорной реакции для реабилитационных или тренировочных упражнений.[2] С помощью исследований, проводимых на стабилметрической платформе, можно выявить такие функциональные нарушения и заболевания, как:

1. нарушения, связанные с заболеваниями позвоночника, нервной и вестибулярной системы;
2. дисфункция голеностопных суставов и опороспособности конечностей;
3. посттравматический и послеоперационный дисбаланс при сколиозах.

II. ОБЗОР УСТРОЙСТВА

Устройство состоит из двух платформ, каждая из которых расположена на четырех чувствительных тензодатчиках. Каждая платформа может работать независимо от другой. Тензодатчики подключены по отдельности, что позволяет обрабатывать данные с каждого из них, позволяя построить картину усилий, добиться большей точности и возможности последующей калибровки.



Рис. 1 – Тензодатчик

Конструкция устройства позволяет добиться более точных данных и отслеживать показание по каждой из конечностей. Исходя из усовершенствований конструкции устройства, улучшается не

только точность измерений, но и появляется возможность по добавлению новых функций и усовершенствованию главной части любого устройства, а именно программного обеспечение.

Устройство предполагает работу в нескольких режимах:

1. определение разности воздействий на каждую из платформ;
2. определение распределения воздействия на каждый тензодатчик;
3. определение стороны платформ, на которую приходится большее воздействие;
4. определение веса пользователя.

Для работы устройства необходимо использовать мобильное приложение, в котором будет возможным переключение режимов работы, просмотр визуально-простых изображений, отображающих результат, а также сохранение результатов предыдущих измерений и их сравнение между собой. Это позволит выявлять не только появление заболевания, но и отслеживать динамику развития дисфункций опорно-двигательного аппарата человека на разных стадиях заболевания.[3–4]

Для данного устройства было разработано программное обеспечение со следующей структурой: в процессе загрузки, пользователь имеет возможность отсканировать QR-код, что позволяет перейти в группу, предоставляющую полезные советы по питанию и рекомендации по упражнениям для поддержания нормальной координации движений. В то же время инициализируются микросхемы HX711. После загрузки, пользователю предлагается нажать на любое место на экране для запуска базовой программы устройства. Базовая программа представляет собой экран, отображающий информацию о весе пользователя. Если пользователь активирует модуль Bluetooth на устройстве (путем нажатия соответствующего значка на экране, который при этом меняет цвет на зеленый), информация о его весе передается по Bluetooth на Android-устройство для построения зависимостей и анализа полученных данных. В Android-приложении пользователь также должен зарегистрироваться, указав свой пол, рост и возраст. Результаты анализа выводятся на экран смартфона.

При рассмотрении основных режимов работы устройства стоит отметить возможность переключения между ними, как при помощи сенсорного экрана самого устройства, так и с использованием передачи соответствующего сигнала по Bluetooth с Android-устройства. Режим считывания показаний с 8 тензодатчиков представляет собой экран с нумерацией датчиков, оформленной в формате "L1, R2 где L1 соответствует верхнему левому датчику на пластине, расположенной слева, а R2 – правому верхнему датчику на пластине справа. Нумерация начинается с левого верхнего

угла и следует по часовой стрелке. Если пользователь включил модуль Bluetooth, данные с датчиков передаются на Android-устройство для последующего анализа.

Следующий режим – считывание данных с одной платформы. Для начала пользователю предлагается встать на одну из платформ. Программное обеспечение автоматически определяет, на какой платформе находится пользователь, и указывает на это, помечая ее как "L" или "R". Дальнейший механизм работы аналогичен режиму считывания данных с 8 тензодатчиков, за исключением того, что в этом режиме на экране устройства визуализируется давление в режиме реального времени. На экране появляется условное изображение платформы с обозначением датчиков по углам. Если нагрузка на один из датчиков значительно выше, чем на других, то он окрашивается в желтый цвет, в противном случае все датчики окрашиваются в синий цвет. Если различий в нагрузке нет или они незначительны, то все датчики окрашиваются в зеленый цвет.[5–7]

III. ЗАКЛЮЧЕНИЕ

В ходе работы было сконструировано изделие – мобильное устройства для диагностики дисфункций опорно-двигательного аппарата человека, действие которого опирается на принципы стабилотрии, свойства тензодатчиков при действии на них сил тяжести. Устройство предназначено для личного использования. Оно призвано помочь следить за состоянием опорно-двигательного аппарата и в случае отклонений обратиться к специалисту.

СПИСОК ИСТОЧНИКОВ

1. Функциональная диагностика [Электронный ресурс] – Режим доступа: <https://posturology.ru/functdiagnost/stabilometriya.php>. –Дата доступа:19.10.2023.
2. Стабилотрическое исследование: краткое руководство / Д. В. Скворцов – М.: Маска, 2010. – 172 с.
3. Биологическая обратная связь по опорной реакции: методология и терапевтические аспекты / О. О. Кубряк [и др.]. –М. : Маска, 2015. – 126 с.
4. Электронные весы на базе HX711 [Электронный ресурс] – Режим доступа: <https://cxem.net/mc/mc416.php>. – Дата доступа: 11.09.2023.
5. Description of STM32F4 HAL and low-layer drivers [Электронный ресурс] – Режим доступа: <https://www.st.com/resource/en/usermanual/um1725description.of.stm32.pdf>. –Дата доступа: 11.09.2023.
6. HX711–STM32–Library [Электронный ресурс] – Режим доступа: <https://github.com/PCov3r/HX711--STM32--Library>. – Дата доступа: 11.09.2023.
7. ILI9341–16bi–FSMC–stm32f407–touch–servo [Электронный ресурс] – <https://github.com/offpic/ILI9341--16bit--FSMC--stm32f407--touch--servo>. – Дата доступа: 11.09.2023.

ОПТИМИЗАЦИЯ ВЫЧИСЛИТЕЛЬНЫХ ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ НА CPU

Соколов А. П.

YADRO

Москва, Российская Федерация

E-mail: andrey.sokolov@yadro.com

В работе рассматривается вопрос оптимизации задач машинного обучения и искусственного интеллекта на процессорах общего назначения (CPU). Перечислены некоторые причины, почему оптимизация данных задач актуальна для CPU. Показаны основные пути оптимизации: оптимизация математических алгоритмов, оптимизация моделей машинного обучения и, наконец, усиление вычислительных возможностей CPU за счет введения в их состав специализированных инструкций. Для каждого из путей показаны конкретные успешные примеры того, как удалось существенно повысить производительность задач машинного обучения на CPU.

ВВЕДЕНИЕ

В настоящее время задачи машинного обучения и искусственного интеллекта требуют все большего объема вычислительных ресурсов. Часто для решения этих задач используют специализированные вычислители. Например, графические карты (GPU). При этом процессоры общего назначения зачастую незаслуженно не рассматриваются. В тоже время есть ряд причин, почему CPU стоит рассматривать как платформу выбора для ряда задач машинного обучения.

Во-первых в CPU используется стандартная модель программирования. Это существенно снижает затраты на разработку и адаптацию алгоритмов под изменяющиеся требования. Во-вторых существует широчайший спектр CPU в смысле энергопотребления (от 10 мВт до 100 Вт при частоте 1 ГГц). Это позволяет более точно подобрать процессор под ту или иную задачу. В-третьих, CPU гораздо более доступны по сравнению со специализированными ускорителями, благодаря тому, что существует множество независимых разработчиков и производителей и архитектур (x86, ARM, RISC-V и др.). В-четвертых, на CPU-системах доступен гораздо больший объем оперативной памяти, чем на специализированных ускорителях. Например, на типовых серверных платформах доступно до 16 ТБ RAM, в то время как на GPU – максимум 128 ГБ. В ряде задач машинного обучения это требование является принципиальным.

Стандартные реализации задач машинного обучения могут обладать неудовлетворительными характеристиками производительности на CPU. Из общих соображений возможны следующие пути оптимизации их быстродействия:

- оптимизация математических алгоритмов, лежащих в основе задач машинного обучения;
- оптимизация используемых моделей;
- усиление возможностей самих CPU, например, за счет специализированных инструкций.

Далее, мы кратко рассмотрим некоторые успешные примеры оптимизации задач машинного обучения в каждом из этих направлений.

I. ОПТИМИЗАЦИЯ МАТЕМАТИЧЕСКИХ АЛГОРИТМОВ

Одним из ярких примеров того, как появление нового математического алгоритма позволяет значительно увеличить скорость решения практически важных задач машинного обучения, является применение метода Винограда для быстрого вычисления сверточных нейросетевых моделей (CNN).

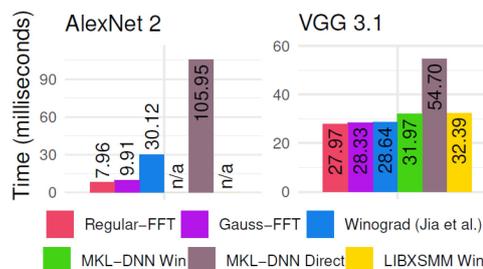


Рис. 1 – Время вычисления CNN моделей при помощи различных алгоритмов [1]

Другим примером эффективности оптимизации математических алгоритмов является фреймворк CatBoost. До появления этого фреймворка для обработки табличных данных в машинном обучении обычно использовались модели градиентного бустинга (GBDT), состоящие из решающих деревьев общего вида. Авторы фреймворка CatBoost предложили использовать симметричные решающие деревья [2]. В итоге модели CatBoost могут быть вычислены значительно быстрее, чем модели на решающих деревьях общего вида. В том числе на CPU.

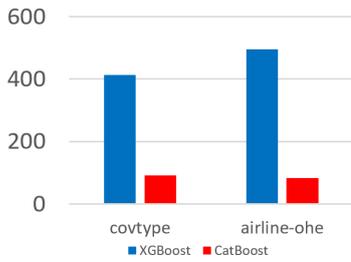


Рис. 2 – Время вычисления GBDT-ансамбля, обеспечивающего точность 95%, мс
Платформа: Intel(R) Xeon(R) Gold 6248R

II. ОПТИМИЗАЦИЯ МОДЕЛЕЙ

Известно, что существует значительная избыточность информации в параметрах глубоких моделей. Вещественные веса нейросетей можно квантовать в числа с низкой точностью (8 бит и менее) или вообще «занулять» (прунинг). При этом, итоговое качество модели практически не снижается. Это явление называется «перепараметризованностью» моделей машинного обучения. Сочетание прунинга и квантизации позволяет «сжимать» глубокие модели в десятки раз практически без потери качества.

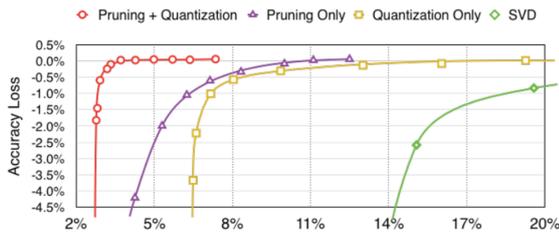


Рис. 3 – Потеря точности глубоких модели в зависимости от метода и степени сжатия [3]

Любопытно, что подобные возможности по существенному сжатию моделей имеются и в семействе моделей градиентного бустинга.

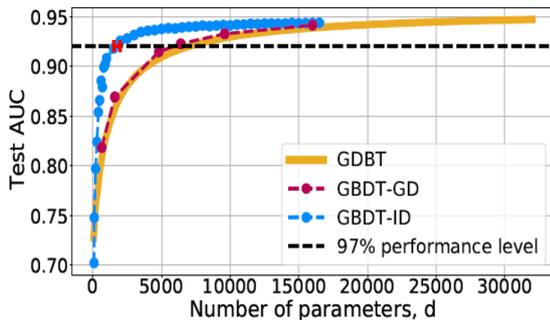


Рис. 4 – Внутренняя размерность GBDT ансамбля (датасет - Epsilon) [4]

III. СПЕЦИАЛИЗИРОВАННЫЕ ИНСТРУКЦИИ

Введение в состав CPU аппаратных блоков, реализующих дополнительные инструкции, является мощным инструментом, позволяющим существенно повысить производительность при решении определенных задач.

Так, например, расширение AVX-512 Core Advanced Matrix Extensions (AMX) вводит в состав процессора специализированный матричный ускоритель и набор инструкций для работы с ним. Данный ускоритель предназначен для более быстрого вычисления операции перемножения матриц, которая является очень востребованной в моделях машинного обучения.

Ниже на рисунке показан пример того, как соотносятся производительности процессоров Intel(R) на задаче детектирования лиц. При этом рассматриваются различные наборы специализированных инструкций, начиная от самых старых (AVX-512 Core) и заканчивая самыми современными (AVX-512 Core AMX).

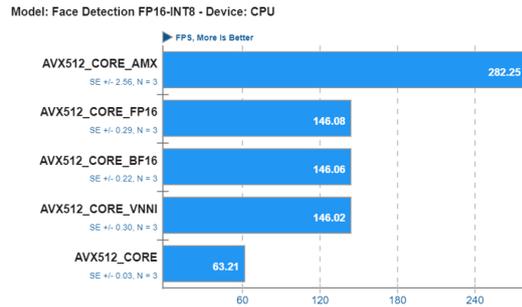


Рис. 5 – Сравнение производительности CPU Intel(R) с различными наборами специализированных инструкций [5]

IV. ЗАКЛЮЧЕНИЕ

Системы на базе CPU доступны, легко программируемы и могут быть быстро адаптированы под новые вычисления. Видно, что современные серверные процессоры успешно адаптируются к возрастающим запросам со стороны задач машинного обучения. Это позволяет им занимать существенную долю серверного сегмента. Оптимизировать задачи машинного обучения на CPU можно разными способами: искать новые и оптимизировать существующие алгоритмы обучения и исполнения моделей, оптимизировать сами модели, а также адаптировать возможности CPU за счет специальных расширений. Все это позволяет значительно сократить разрыв между CPU и специализированными ускорителями во многих практически-важных случаях.

1. Zlateski A., Zhen J., Kai L. and Fredo D., "The anatomy of efficient FFT and winograd convolutions on modern CPUs. Proceedings of the ACM International Conference on Supercomputing, 2019, pp. 414-424. doi:10.1145/3330345.3330382
2. Prokhorenkova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A., "CatBoost: unbiased boosting with categorical features NeurIPS, 2018.
3. Han S., Mao H., Dally W.J., "Deep Compression: Compressing Deep Neural Network with Pruning, Trained Quantization and Huffman Coding ICLR, 2016.
4. Ivanov S., Litovchenko L., Prokhorenkova L., Sokolov A.P., "Tree Ensembles with Gradient Descent: Loss Landscape, Intrinsic Dimension and Split Permutation Invariance OpenTalks.AI, 2023.
5. <https://www.phoronix.com/review/intel-xeon-amx>

АНАЛИЗ АЛГОРИТМОВ ПОИСКА ПУТИ В НАВИГАЦИОННЫХ СЕТКАХ

Хирьянов И. Д.

Факультет информационных технологий,
Полоцкий государственный университет имени Ефросинии Полоцкой
Полоцк, Республика Беларусь
E-mail: i.khiryanaou@psu.by

Проанализированы алгоритмы поиска пути в навигационных сетках, взяты одни из самых распространенных: поиск в ширину, Дейкстры и A^ . Для тестирования использован сервис PathFinding и выделены ключевые критерии: длина пути, время и количество итераций. В ходе исследования выявлено что A^* находит кратчайший путь как за меньшее время, так и совершает минимальное количество операций по сравнению с другими рассмотренными алгоритмами.*

ВВЕДЕНИЕ

Выбор оптимального пути для каждого объекта является одной из самой важных задач искусственного интеллекта в игровой индустрии. Он описывается как процесс определения набора перемещений объекта из одной точки в другую, без столкновения с любыми препятствиями. Оптимальный путь должен обладать тремя свойствами: длиной пути, временем и количеством итераций для поиска.

Первое свойство определяет расстояние, требуемое для прохождения пути. При использовании этого показателя, оптимальный маршрут является кратчайшим. Это означает, что дистанция от начала движения до финиша не длиннее любого другого пути.

Второе свойство определяет оптимальный путь, как самый быстрый. Это означает, что время, затраченное на прохождение маршрута, всегда меньше, чем при прохождении любого другого.

Третье свойство определяет сколько ресурсов вычислительной системы потребуется для нахождения оптимального маршрута. Соответственно наименьшее количество итераций является лучшим показателем среди сравнимых алгоритмов.

Требуется проанализировать самые распространенные алгоритмы для нахождения пути и выбрать самый эффективный из них.

I. АЛГОРИТМ ПОИСКА В ШИРИНУ

Одним из самых первых алгоритмов поиска кратчайшего пути на графах является алгоритм поиска в ширину [1-2].

Алгоритм:

1. Выбираем начальную вершину, задаём ей нулевой уровень глубины.
2. Всех её соседей добавляем в очередь для посещения, а саму помечаем как посещённую и больше к ней не обращаемся.
3. Берём элементы из очереди для посещения, считая их уровень глубины, как: уровень

- глубины родителя + 1. Добавляем эти соседние вершины в очередь для посещения.
4. Текущую вершину, помечаем как посещённую и удаляем из очереди для посещения.
 5. Повторяем шаги 3 и 4, пока не обойдём весь граф.

II. АЛГОРИТМ ДЕЙКСТРЫ

Алгоритм Дейкстры работает только для графов без рёбер отрицательного веса [3-4].

Алгоритм:

1. Задаём набор расстояний, в котором будут храниться кратчайшие расстояния от исходной вершины до всех остальных вершин графа, а также набор посещённых вершин.
2. Начальной вершине задаём значение равное нулю, а всем остальным бесконечности.
3. Пока в набор посещённых вершин не добавлены все вершины делаем следующее: выберем вершину, не включённую в наш набор посещённых вершин, значение которой минимально; добавим её в наш набор посещённых вершин; обновим значения всех смежных вершин, как: минимальное значение суммы текущей вершины и веса ребра.

III. АЛГОРИТМ A^*

Алгоритм A^* является усовершенствованием алгоритма Дейкстры за счёт введения эвристики. Переход осуществляется в ту вершину, предположительный путь от которой до конечной является кратчайшим. Чаще всего пользуются эвристикой Манхеттена, но помимо неё возможно использование эвристик Чебышева и Евклида [5].

Алгоритм:

1. Добавляем начальный узел в открытый список для хранения узлов
2. Повторяем следующее: Останавливаемся, если добавили конечный узел в открытый список, в этом случае путь найден. Или открытый список пуст, и мы не дошли до конечного узла. Здесь получаем, что не существует пути до конечного узла. Ищем в открытом списке узел с наименьшим значением целевой функции F . Делаем его текущим узлом.

Помещаем этот узел в закрытый список. (И удаляем с открытого) Для каждого из соседних узлов делаем следующее: если узел находится в закрытом списке, игнорируем ее. В противном случае если узел еще не в открытом списке, то добавляем его туда. Делаем текущий узел родительским для этой клетки. Рассчитываем стоимости F, G и H узла. Если узел уже в открытом списке, то проверяем, не дешевле ли будет путь через этот узел. Для сравнения используем стоимость G. Более низкая стоимость G указывает на то, что путь будет дешевле. Если это так, то меняем родителя узла на текущий узел и пересчитываем для него стоимости G и F.

3. Сохраняем путь. Двигаясь назад от целевого узла, проходя от каждого узла к его родителю до тех пор, пока не дойдем до начального узла. Это и будет наш путь.

IV. СРАВНЕНИЕ АЛГОРИТМОВ ПОИСКА

Для проведения тестов и выявления лучшего из 3-х представленных алгоритмов была проведена работа с сервисом <https://qiao.github.io/PathFinding.js/visual/>, в котором реализованы в единой оболочке большинство алгоритмов поиска путей. В качестве критериев оценивания были выбраны: скорость нахождения кратчайшего пути, количество итераций, выполненных в процессе поиска, и длина найденного пути. Т.к. алгоритмы реализованы на одной и той же платформе (java script), то в расчёт не берётся скорость компилятора.

На рисунке 1 продемонстрирован первый тест. В этом и последующих зеленый квадрат - начало пути, красный - конец, черный - непроходимое препятствие, а все остальное это поле по которому можно передвигаться. Результаты тестирования описаны в таблице 1.

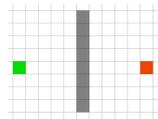


Рис. 1 – Первый тест

Таблица 1 – Результаты первого теста

Алгоритм	Длина пути	Время(ms)	Кол-во операций
A*	13,31	1,1	136
Дейкстра	13,31	1,8	1059
BFS	13,31	1,4	843

Исходя из первого тестирования видно, что оптимальный маршрут по длине находят все. Дейкстра тратит на это больше всего времени и количество операций. Алгоритм A* показывает наилучший результат по всем критериям.

На рисунке 2 показан второй тест, а его результаты - в таблице 2.

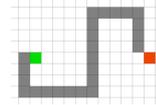


Рис. 2 – Второй тест

Таблица 2 – Результаты второго теста

Алгоритм	Длина пути	Время(ms)	Кол-во операций
A*	16,49	0,2	133
Дейкстра	16,49	2,0	1267
BFS	16,49	1,0	1420

Второе тестирование показало, что A* все еще превосходит другие алгоритмы. Дейкстра в этот раз показал себя лучше BFS по количеству операций, хотя все еще уступает во времени нахождения.

На рисунке 3 продемонстрирован третий тест. В таблице 3 можно ознакомиться с его результатами.

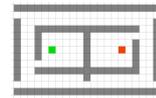


Рис. 3 – Третий тест

Таблица 3 – Результаты третьего теста

Алгоритм	Длина пути	Время(ms)	Кол. операций
A*	29,31	0,3	320
Дейкстра	29,31	9,3	2458
BFS	29,31	9,0	2542

Третье тестирование показало те же результаты, что и второе. Алгоритм A* остается лучшим.

Из таблицы проведенных тестов видно, что A* находит кратчайший путь как за меньшее время, так и совершает минимальное количество операций по сравнению с другими рассмотренными алгоритмами.

СПИСОК ЛИТЕРАТУРЫ

1. Л.А. Павлов, Н.В. Первова Глава 6 Структуры и алгоритмы обработки данных // Алгоритмы на графах / Издательство «Лань», 2020г. – 199 – 218 с.
2. Томас Кормен, Чарльз Лейзерсон, Рональд Ривест, Клиффорд Штайн Глава 22.2 Поиск в ширину // Алгоритмы. Построение и анализ второе издание/ Издательский дом "Вильямс 2011г. - 613-622 с.
3. Ананий В. Левитин Глава 9. Жадные методы: Алгоритм Дейкстры // Алгоритмы: введение в разработку и анализ = Introduction to The Design and Analysis of Algorithms. - Издательский дом "Вильямс 2006г. - 189-195 с.
4. Адитья Бхаргава Глава 7 Алгоритм Дейкстры // Грокаем алгоритмы. Иллюстрированное пособие для программистов и любопытствующих / Издательство "Питер 2017г. - 151-181с
5. С.Рассел, П. Норвиг Глава 3 Решение задач по средствам поиска // Искусственный интеллект. Современный подход. Четвертое издание. Том I. Решение проблем: знания и рассуждения / Издательство "Диалектика 2021г. - 125 -204 с.

РЕДАКТОР ТЕКСТОВ СТРУКТУРИРОВАННОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Хорошавин В. Д., Насевич П. Е., Садовский М. Е.
Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {vasya.khoroshavin, fallenchromium}@gmail.com, sadovski@bsuir.by

В данной работе на основании сравнения существующих программных средств для представления и редактирования знаний выдвинуты требования к редактору текстов структурированного представления знаний интеллектуальных систем, а также предложен подход к его построению на основании языка унифицированного смыслового представления знаний и открытой семантической технологии проектирования интеллектуальных систем.

На сегодняшний день информационные технологии играют неотъемлемую роль в жизни людей: они используются для автоматизации рутинных, помощи в трудных интеллектуальных задачах, обеспечения совместной работы и коммуникации. Информационные компьютерные технологии широко применяются в научной среде для сохранения и обработки информации с целью управления знаниями. Сюда также можно отнести каталогизацию и генерации новых знаний, обрательное их в научные труды, чтобы другие люди могли получить доступ и эффективно ими воспользоваться. Такое использование информационных технологий предоставляет широкие возможности и имеет большие перспективы, но, тем не менее, обладает рядом недостатков. На данный момент одна из ключевых проблем заключается в инструментарии, который служит для обеспечения сохранения и редактирования знаний. Он не является оптимальным по нескольким параметрам:

- простота переноса знаний из нецифровых источников в компьютерную систему;
- возможность переиспользования записанных знаний в статьях или иных научных трудах;
- автоматизация генерации, визуализации и прочей обработки знаний;
- представление знаний в хорошо воспринимаемом человеком формате;
- эффективность и интуитивная понятность пользовательского интерфейса;
- возможности для совместной работы и коллаборации.

Многие из этих проблем в той или иной степени решены в разных инструментах и компьютерных системах, однако на данный момент проблематично найти такую систему, которая решала бы все описанные выше задачи. В данной работе предлагается совместить один из современных подходов к редактированию текстовой информации (т.н. «блочное редактирование») с онтологическим подходом к структурированию знаний и

языком унифицированного смыслового представления знаний (SC-кодом) [1]. Так пользовательский интерфейс будет спроектирован посредством формализации предметной области компонентов пользовательского интерфейса, предметной области действий пользователя и предметной области внешних информационных конструкций и языков и разработкой решателя задач для пользовательского интерфейса [2]. Такой подход позволит реализовать концепцию адаптивного интерфейса, который будет подстраиваться под навыки и нужды пользователя, а также позволит без больших затрат времени интегрировать компоненты интерфейса, которые будут реализованы позже.

Был проведён сравнительный анализ редакторов исходных текстов баз знаний и редакторов текста в целом, и среди альтернатив были выделены следующие кандидаты: Overleaf [3], Notion [4], КВЕ [5], Protege [6]. По итогу анализа был выдвинут список требований к новому редактору баз знаний:

- Для обеспечения беспрепятственного ввода формализованных знаний необходим продуманный, эргономичный и задокументированный пользовательский интерфейс. Функции такого интерфейса включают в себя:
 - возможность использования горячих клавиш (в т.ч. возможности настроить персональные горячие клавиши);
 - графические методы редактирования текста (перетаскивание, контекстные меню, палетка команд и т.д.);
 - богатая обратная связь от компьютерной системы (реализация подсказок по наведению, документации и интерактивных инструкций по работе с программой);
 - активная помощь от компьютерной системы (реорганизация интерфейса в зависимости от текущих нужд пользователя или режима редактирования, проработанные сообщения об ошибках, механизм автодополнения текста).

– Для обеспечения возможности переиспользования записанных знаний в статьях или иных научных трудах необходим язык представления знаний, над которым определены операции перевода этого текста в требуемый формат и визуализацию. В случае описания онтологии компонентов пользовательского интерфейса редактора баз знаний на этом же языке представления знаний, появляется возможность реализовать реорганизацию интерфейса и механизмы помощи пользователю более оптимальным образом как с точки зрения программиста, так и с точки зрения пользователя. В качестве языка представления знаний предлагается использовать язык унифицированного представления знаний, SC-код. Это решение обеспечит следующие преимущества:

- Машина обработки знаний, при добавлении в неё соответствующих правил, сможет давать подсказки высокого качества и предлагать пользователю системы более разнообразные и персонализированные алгоритмы помощи в написании научных статей.
- Описание пользовательского интерфейса на языке унифицированного представления знаний позволит более просто менять техническую реализацию интерфейса и его свойств, т.к. для этого достаточно будет лишь описать внешний язык представления интерфейса и правила представления компонентов интерфейса на этом внешнем языке.
- Выбор SC-кода в качестве языка представления знаний даёт возможность интегрировать редактор с любыми *ostis*-системами, т.е. расширение функционала редактора станет доступной задачей даже для людей не знакомых с программированием. Сам же редактор можно будет использовать в составе более масштабных *ostis*-систем.

Помимо прочего, использование SC-кода позволит решить задачу автоматизации генерации, визуализации и другой обработки знаний, т.к. машина обработки знаний позволяет использовать поиск по графовым шаблонам, использовать подсистему логического вывода, в т.ч. по настраиваемым правилам, которые представлены на том же языке что и знания, над которыми проводятся операции. Таким образом операционная нагрузка

на научных работников будет уменьшена, т.к. необходимые операции обработки и вывода знаний нужно будет описать лишь единожды.

Представление знаний в хорошо воспринимаемом человеком виде и эффективность, понятность пользовательского интерфейса будут ключевыми целями при разработке пользовательского интерфейса. Реализации этих требований будет способствовать использование концепции блочного редактирования текста. Эта концепция стала широко применяться в редакторах текстов, которые были разработаны недавно, из-за того, что она предоставляет широкие и интуитивно-понятные возможности для реорганизации текста и выражения отношений между параграфами. Суть концепции заключается в том, что каждая обособленная единица информации выделяется в отдельный блок, над которым можно производить преобразования – поменять порядок следования блоков, сделать блок подблоком другого блока, вставить ссылку или отображение блока в другой блок. Для онтологического подхода к структурированию знаний наиболее важной является возможность выражать отношения между блоками с помощью концепции «подблок блока». В сочетании с SC-кодом, в котором заданы различные виды отношений между сущностями и имеется возможность описать пользовательские отношения для заданной предметной области, блочное редактирование должно стать мощным инструментом при редактировании научных статей или онтологий. Из всего вышесказанного следует, что разработка подобного редактора принесёт большой прирост в эффективности работы над трудными интеллектуальными задачами, а потому разработка целесообразна и перспективна.

СПИСОК ЛИТЕРАТУРЫ

1. Технология комплексной поддержки жизненного цикла семантически совместимых интеллектуальных компьютерных систем нового поколения / Под ред. В. В. Голенкова. – Минск: Бестпринт, 2023. – 1064 с.
2. Открытые семантические технологии проектирования интеллектуальных систем : сб. научн. тр. / Белор. гос. ун-т информ. и радиоэлектрон. ; редкол.: В. В. Голенков (отв. ред.) [и др.]. – Минск. : Бестпринт, 2021. – 330 с.
3. Overleaf documentation [Электронный ресурс]. – Режим доступа: <https://www.overleaf.com/learn>. – Дата доступа: 26.10.2023.
4. Notion. – Режим доступа: <https://www.notion.so/>. – Дата доступа: 26.10.2023.
5. KBE, Knowledge Base source Editor. – Режим доступа: <https://github.com/ostis-ai/kbe>. – Дата доступа: 26.10.2023.
6. Protege. – Режим доступа: <https://protege.stanford.edu/>. – Дата доступа: 26.10.2023.

АНАТОМИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Черепанова Т. В., Корняков К. В.
Департамент разработки ПО искусственного интеллекта
YADRO
Нижний Новгород, Россия
E-mail: t.cherepanova@yadro.com, kirill.korniakov@yadro.com

Системы искусственного интеллекта – крайне обширная область знаний, стремительно развивающаяся в настоящее время, и на наших глазах меняющая мир вокруг нас, проникая во все технологические области. В ходе выступления мы попробуем заглянуть «под капот» систем, использующих искусственный интеллект, разобратся из каких программных и аппаратных компонент они собираются, и с какими сложностями сталкиваются их разработчики. Также, рассмотрим основные тренды в технологиях искусственного интеллекта, и какую роль в будущем индустрии играют производители платформ.

ВВЕДЕНИЕ

Посмотрим на определение искусственного интеллекта – это область компьютерных наук, которая занимается разработкой и созданием систем и программ, способных имитировать когнитивные способности человека, а именно, анализ информации, обучение из опыта, решение задач и принятие решений. Искусственный интеллект может включать в себя такие методы и технологии, как машинное обучение, нейронные сети, обработка естественного языка, компьютерное зрение, робототехника и другие. Он может применяться в различных сферах, включая медицину, робототехнику, экономику, финансы, игровую индустрию, автомобильную промышленность и другие.

Искусственный интеллект имеет потенциал изменить мир и повлиять на многие аспекты нашей жизни, повысив эффективность и точность многих процессов. Однако существуют и некоторые этические и социальные вопросы, связанные с применением искусственного интеллекта, которые требуют внимательного рассмотрения и регулирования. Искусственный интеллект, может быть, и источником определенных опасностей в случае неправильного обучения или использования, или привести к развитию зависимости, которая может привести к утрате человеческих навыков и способностей.

Несмотря на широкую распространённость технологий сегодня, область искусственного интеллекта продолжает активно развиваться, появляются новые модели использования, создаются все более сложные алгоритмы, которые в свою очередь требуют более мощных вычислительных ресурсов.

I. ОБЛАСТИ ПРИМЕНЕНИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Искусственный интеллект имеет широкий спектр применений в различных областях. Некоторые из них включают:

- Медицина: ИИ используется для диагностики и прогнозирования заболеваний, создания индивидуализированных лечебных пла-

нов и разработки новых лекарственных препаратов.

- Финансы: ИИ применяется для анализа больших объемов финансовых данных, определения трендов и паттернов, прогнозирования рынка и принятия инвестиционных решений.
- Производство: ИИ помогает автоматизировать процессы производства, решать проблемы сборки и контроля качества, оптимизировать время и ресурсы.
- Коммуникации: ИИ используется для разработки систем автоматического распознавания речи, машинного перевода, создания виртуальных помощников и ботов для общения с клиентами.
- Безопасность: ИИ помогает обнаруживать аномальное поведение в системах безопасности, распознавать лица и голоса, анализировать видео материалы для предотвращения преступлений.

Это лишь некоторые области, применение искусственного интеллекта уже стало стандартом де-факто, уже изменило нашу жизнь, и потенциал развития огромный, и его применение будет продолжать расти и проникать во все области информационных технологий, сферы жизни и бизнеса. Предметная область носит стремительный, подрывной характер технологических инноваций, и открытые подходы и стандарты поддерживают это масштабное развитие. Например, 6 лет назад не было активного распространения технологий глубокого обучения и нейронных сетей, а использовали традиционное компьютерное зрение.

II. ЛАНДШАФТ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Для создания решений с использованием искусственного интеллекта требуются не только алгоритмические библиотеки и модели, но и множество технологий вокруг, обеспечивающих работу с данными, эффективное использование вычислительных ресурсов и операционное управление. Группа аналитиков регулярно выпускает отчет Machine Learning, Artificial Intelligence and

Data (MAD) Landscape, где в виде таблицы-карты собраны все ключевые компании и технологии, которые функционируют в области искусственного интеллекта, таблица огромна и стремительно меняется. Ландшафт включает в себя следующие ключевые направления:

- Инфраструктура: Различные системы хранения, доступа и защиты данных, оркестрация, мониторинг, системы безопасности.
- Аналитика: Платформы анализа и визуализации данных
- Машинное обучение и искусственный интеллект: универсальные инструменты для разработки, тестирования и развертывания моделей, подготовки данных и специализированные решения для определённых технологий.
- Приложения: множество готовых решения для широкого спектра задач.

Большинство решений базируется на фундаменте открытых (open-source) библиотек, моделей, технологий и стандартов, и тем самым открытый подход поддерживает бурный рост и развитие.

Первые шаги в создании искусственного интеллекта были сделаны еще в 1940-1950 годах. В это время ученые искали способы моделирования человеческого интеллекта в машинах. Одним из ранних прорывов стало создание теста Тьюринга в 1950 году, разработанного Аланом Тьюрингом. Активное развитие технологий началось в 1990-ые годы, когда появились более мощные компьютеры. В 1997 году компьютер Deep Blue созданный IBM, победил чемпиона мира по шахматам Гарри Каспарова. Рассмотрим стремительность эволюционного развития на примере языковых моделей. Первым широкоиспользуемым алгоритмом, основанным на принципах языковой модели, в начале 2010-х стал продвинутый «Т9», работающих на телефоне. В средах разработки появились подсказки ввода. В 2013 году появилось векторное представление слова. В 2018 году появилась первая GPT (Generative Pre-trained Transformer) модель, способная генерировать текст. GPT-2 состояла из 1.5 миллиарда параметров и 40 Gb данных тренировки и считается канонической моделью на которой основано дальнейшее развитие. Сегодня доступны открытые модели размером до 100 миллиардов параметров, а количество разных моделей на популярном ресурсе Hugging Face, поддерживающих диалог, превышает 2500.

III. СЛОЖНОСТИ ПОСТРОЕНИЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Фундаментальные принципы создания систем с использованием технологий искусственного интеллекта опираются на алгоритмы (или модели) и качественные входные данные, на которых эта модель будет обучена. Также немаловажным является интеграция искусственной

системы в другие технологические процессы, обеспечивая качественную работу и эффективное доверительное взаимодействие с человеком. Если заглянуть внутрь технологического стека, то его можно условно разделить на три уровня:

1. приложения и задачи, которые приносят непосредственно пользу;
2. программное обеспечение среднего слоя (framework или middleware) где реализованы сами алгоритмы;
3. аппаратное обеспечение.

Все три составляющих подбираются под конкретную задачу, объем данных и доступные вычислительные ресурсы. Рассмотрим три основные проблемы, которые возникают внутри технологического стека, которые могут быть не видны разработкам конечных приложений, но над их решением трудятся тысячи инженеров.

Проблема переносимости ПО и фрагментация

Существует несколько вычислительных аппаратных платформ доступных на рынке от разных производителей и каждый предлагает свое собственное решение, включая программный стек. Приложение разработанное под одну платформу не будет работать на другой, возможности переноса ограничены. Некоторые библиотеки и фреймворки могут не доступны на некоторых платформах. Наблюдаются тренды на унификацию и стандартизацию начиная от языков программирования до API алгоритмических библиотек. При разработке решения рекомендуется внимательно анализировать доступные библиотеки, опираться на уровень абстракции, когда переносимость будет бесшовной или с минимальными изменениями.

Комбинаторный взрыв вариантов реализации

Обеспечения функциональной переносимости в большинстве случаев будет недостаточно. Как правило алгоритмы искусственного интеллекта требуют значительных вычислительных ресурсов для исполнения и обучения. Внутри реализации алгоритма необходима оптимизация под определенную платформу, и здесь количество вариаций многократно увеличивается. Для уменьшения количества вариаций алгоритмы реализуются на языках, поддерживающих параллельные вычисления, таких как OpenCL, CUDA или SYCL и производители платформ предоставляют современные компиляторы основанные на LLVM технологии с IR представлением. К сожалению, не всегда можно достичь максимальной производительности с языками и компиляторные подходы выходят на доменный уровень, такие технологии как MLIR и TVM, где оптимизируется операции

высокого уровня и как кубики собираются в вычислительный конвейер. Закрытые экосистемы производителей накладывают ограничения на оптимизации, по сути их приходится разрабатывать отдельно под каждую экосистему. Современный тренд на открытые аппаратные платформы такие как RISC-V будет драйвером объединения экосистем и ускорит развитие технологий в целом.

Большие данные

Объем данных непрерывно растет и вычислительный ресурс одной машины оказывается недостаточными для решения множества задач искусственного интеллекта. Существуют подходы и библиотеки распределённых вычислений как низкого уровня – MPI, так и уровня приложения – Spark, которые предоставляют возможность выполнения одной задачи на нескольких узлах одновременно. Однако, использование данных методов усложняет процесс не только алгоритмического программирования, но и требует заблаговременного обдумывания масштабов данных и нагрузок на вычислительные ресурсы еще на стадии проектирования системы.

IV. ТРЕНДЫ В ТЕХНОЛОГИЯХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Технологии продолжают свое неуклонное развитие, которое описывается циклом зрелости (Hype cycle) предложенным компанией Gartner. На это кривой видно какие из технологий уже подтвердили свою эффективность и активно используются, какие на стадии активного анализа их возможностей, а какие только исследуются. Сегодня в зоне внимания исследователей, экспертов и бизнеса находятся генеративные модели искусственного интеллекта, которые вышли на пик ожиданий. Вслед за ними выстраивается еще десяток захватывающих технологий, такие как

умные роботы, нейроморфные вычисления, ответственный искусственный интеллект. Каждая из них обладает своим уникальным потенциалом и обещает принести революцию в нашу жизнь. Таким образом, научное сообщество стоит на пороге новых достижений в области искусственного интеллекта, которые предоставят не только новые возможности, но и вызовы, требующие от нас как ученых и разработчиков сотрудничества и ответственного использования этих прогрессивных технологий.

V. Выводы

Непрерывно развивающиеся технологии искусственного интеллекта предоставляют и будут предоставлять научный и практический интерес в нескольких направлениях:

- Применение новейших технологий в практической деятельности.
- Исследования новых алгоритмов и подходов в области искусственного интеллекта направлены на поиск эффективных способов решения сложных проблем и улучшения качества жизни людей.
- Погружение в технологические стеки и решение фундаментальных проблем системного и аппаратного уровня, что позволит создавать более мощные и эффективные интеллектуальные системы, способные обрабатывать и анализировать огромные объемы данных.

СПИСОК ЛИТЕРАТУРЫ

1. The 2023 MAD (ML/AI/Data) Landscape [Электронный ресурс] / – Режим доступа: <https://mad.firstmark.com/>. – Дата доступа: 18.10.2023.
2. What's New in Artificial Intelligence from the 2023 Gartner Hype Cycle [Электронный ресурс] / – Режим доступа: <https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2023-gartner-hype-cycle/>. – Дата доступа: 25.10.2023.

ABOUT ENHANCED ACCESS CONTROL USING FIDO2 AUTHENTICATION AND ATTRIBUTES

Zhidovich A., Lubenko A., Vojteshenko I.

Faculty of Applied Mathematics and Computer Science, Belarusian State University

Minsk, Belarus

E-mail: {anton.zhidovich, alexeilubenko02}@gmail.com, voit@bsu.by

In this paper a joint use of the FIDO2 specification and attribute-based access control with support for reading attributes from electronic documents is proposed. Existing solutions and implementation options are reviewed. This provides an opportunity to improve access control methods to information and resources.

INTRODUCTION

When building a corporate system, it's necessary to give special priority to the issue of access to system resources and the differentiation of user rights. The key concepts here are authentication and authorization. These two processes are closely inter-related and must be implemented to meet the security and scalability requirements of the information system, while ensuring usability and accessibility for users.

There are various authentication standards, many of which may not provide a high level of security and may be compatible only with a particular class of software or may be proprietary.

I. RELEVANCE OF FIDO2 TECHNOLOGY

Passwordless authentication allows a user to access an information system without entering a password or answering security questions. Instead, the user provides other form of evidence such as a fingerprint, NFC-chip, or hardware token code [1].

The modern approach to passwordless authentication is the open standard FIDO2, jointly developed by the FIDO Alliance and the W3C consortium. The FIDO2 specification uses public-key cryptography and consists of two groups of standards. One of these is called the W3C WebAuthn standard. The second part is the Client to Authenticator Protocol (CTAP, CTAP1, CTAP2).

The following benefits can be pointed out:

1. Resistance to phishing, MITM, replay attacks;
2. Variety of built-in authenticators with fast secure access;
3. Ease of implementation;
4. Enhanced privacy for users.

See, for example, [2] for a more detailed overview of the FIDO2 specification, its benefits and workflow.

FIDO2 authentication is increasingly used in information systems in different areas. For example, online banking services of such banks as Priorbank, Belinvestbank, Sberbank, and e-mail services like Mail.ru.

II. ATTRIBUTE-BASED ACCESS CONTROL

Attribute-Based Access Control (ABAC) is an authorization model that uses attributes to deter-

mine user rights. The key advantage of ABAC is its flexibility. The method can solve a wide variety of access issues with minimal administrative control [3].

ABAC offers a series of benefits, such as:

1. Possibility to construct similar to business concepts rules;
2. No limitation on the complexity of the rules;
3. Rules with dynamic parameters supported;
4. Possibility of filtering the data user has access to.

This authentication model is getting more demanded and finds application in many fields. In [4], the ABAC model is proposed to be used to manage access to emergency patient data. With the designed prototype, the authors evaluated the resulting authorisation system. ABAC is also used in IAM services on such platforms as Amazon Web Services, Microsoft Azure and Okta.

In the ABAC systems anonymous attributes can be used for access decisions. For example, confirmation of age without the need for personal identification.

The joint use of FIDO2 authentication and ABAC can significantly increase the level of flexibility and scalability of a security system. Importantly, these are two different processes, and the FIDO2 specification (specifically WebAuthn) doesn't provide such integration. Consequently, configuring ABAC and FIDO2 together may require additional effort and resources specific to each system. For example, if dynamic attributes (such as location) are in the system, integration with FIDO2 authentication will require the implementation of mechanisms to collect and update this data without violating the requirements of the specification.

III. ELECTRONIC IDENTIFICATION

The key issue is selecting the appropriate source of user attributes. One suitable option is to use electronic documents that store the issuer's signature.

Electronic identification (eID) is a smart card or biometric passport with an embedded RFID-microchip that contains a biometric identification tool with the holder's data in accordance with the ICAO standard [5]. The following security protocols ensure high level of protection and data integrity:

1. Basic Access Control (BAC) is designed to ensure that card data is only accessed when the card is physically reached.
2. Passive authentication allows the reader to verify eID data authenticity.
3. Chip authentication is designed to protect document data from being altered and cloned.

The use of eID promotes interoperability at the legal, semantic and technical levels, making it an efficient and convenient way of confirming identity in the digital environment.

IV. EXISTING SOLUTIONS

The use of the ABAC model based on OAuth2.0 protocol is proposed in [6]. This is a reasonable solution, but in this case there is no way to verify the user data given by the service provider.

In [7], a complete and industry-ready solution for using anonymous credentials with local or remote attestation through a FIDO2 channel is proposed in the form of the FIDO-AC framework, which is an extension of the basic FIDO2 specification. An evaluation of the security and privacy provided by the system and the realisation of a working prototype are also presented.

The essence of the resulting solution is the creation of an additional trusted party - a mediator. This party is responsible for validating the user's data, for example from an eID document.

Thus, the FIDO-AC identifies the following parties to the authorisation process: FIDO-server (relying party), authenticator, client, mediator, FIDO-AC application and eID. The authorisation workflow largely follows the standard FIDO2 authentication process, which is also described in [2]. The FIDO-AC application is designed to perform BAC and to read data from the user's eID. The mediator, in its turn, should check the verification of the data received from the eID, i.e. perform passive and chip authentication. The FIDO-AC access control process is presented in detail in [7].

V. IMPLEMENTATION OPTIONS

In this section, the use of biometric documents of the Belarus compliant with the ICAO standard is proposed to improve the access control system and ensure user security.

To implement a FIDO server, it's reasonable to use the cross-platform ASP.NET Core. It enables the development of a single solution for both user interface and the web API, and provides extensive deployment and scalability options. To develop a controller that allows the web application to act as a relying party (i.e form FIDO2-compliant queries and verify responses), the component released by Rock Solid Knowledge for the ASP.NET Core 6.0 platform can be used [8]. Web hosting services such as Somee

and Google Cloud Platform are suggested for testing and debugging the web application. Also, most deployment scenarios assume containerisation with Docker. For real use, the application is deployed on the company's corporate server.

Currently, desktop and mobile web browsers have built-in WebAuthn API support, so any of them can act as a client. To generate client-side requests and invoke WebAuthn API, the server will send a JS-program.

Since the FIDO-AC protocol does not require physical separation of the mediator and FIDO-AC application sides, they can be merged into a single application. In this case, no concern for the secure connection between the mediator and the FIDO-AC application is required.

Android can be chosen as the platform for the application, in this case the device must have a built-in NFC-sensor to read data from the eID.

In order to realize correct and proper interaction of android device with biometric document, JMRTD-component [9] can be used. JMRTD provides the ability to connect to an eID, retrieve data from it and perform the verification described above, supporting ICAO standard documents.

REFERENCES

1. Passwordless Authentication [Electronic resource] – Mode of access: <https://www.cyberark.com/what-is/passwordless-authentication/>. – Date of access: 05.10.2023
2. Semantic Approach to Designing Applications with Passwordless Authentication According to the FIDO2 Specification / A. Zhidovich [et al.]. // Open Semantic Technologies for Intelligent Systems – 2023. – № 7. – P. 311-316.
3. Casey K. What is Attribute-Based Access Control (ABAC)? [Electronic resource] / K. Casey // Okta company blog. – Mode of access: <https://www.okta.com/blog/2020/09/attribute-based-access-control-abac/>. – Date of access: 04.10.2023.
4. AC-ABAC: Attribute-based access control for electronic medical records during acute care / Marcela T. de Oliveira [et al.]. // Expert Systems with Applications. – 2023. – Vol. 213, part C.
5. Doc 9303 [Electronic resource] // ICAO. – Mode of access: <https://www.icao.int/publications/pages/publication.aspx?docnum=9303>. – Date of access: 14.10.2023.
6. Belovodov, A. Using attribute-based access control in the OAuth 2.0 protocol / A. Belovodov, O. Laponina // International Journal of Open Information Technologies. – 2023. – Vol. 11, № 6. – P. 182-189.
7. Fast Identity Online with Anonymous Credentials (FIDO-AC) / Wei-Zhu Yeoh [et al.]. // Proceedings of the 32nd USENIX Security Symposium. – 2023 –P. 3029–3046.
8. RSK IdentityServer [Electronic resource] – Mode of access: <https://www.identityserver.com/> – Date of access: 10.10.2023.
9. JMRTD [Electronic resource] – Mode of access: <https://jmrtd.org/>. – Date of access: 11.10.2023.

Секция «Проектирование
встраиваемых систем»

Section «Embedded systems design»

СХЕМНАЯ РЕАЛИЗАЦИЯ В ЗАКАЗНЫХ СБИС РАЗРЕЖЕННЫХ СИСТЕМ ДИЗЬЮНКТИВНЫХ НОРМАЛЬНЫХ ФОРМ БУЛЕВЫХ ФУНКЦИЙ

Бибило П. Н., Кардаш С. Н.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: bibilo@newman.bas-net.by, kardash77@gmail.com

Приводятся результаты экспериментального исследования эффективности программ блочного покрытия и минимизации многоуровневых представлений булевых функций, выполняемых при синтезе комбинационных схем в библиотеке проектирования заказных СБИС (сверхбольших интегральных схем) по системам разреженных ДНФ (дизьюнктивных нормальных форм) булевых функций.

ВВЕДЕНИЕ

Проблема эффективной схемной реализации цифровых комбинационных блоков в заказных СБИС по-прежнему является актуальной при создании средств автоматизированного проектирования цифровых систем. Важным аспектом этой проблемы является то, что современные синтезаторы логических схем чувствительны к форме задания проектной информации, в качестве которой выступают VHDL либо Verilog описания моделей функционирования комбинационных схем – те или иные формы задания систем полностью определенных булевых функций. Синтез логических схем выполняется в два этапа – технологически независимая оптимизация представлений систем булевых функций (этап 1) и технологическое отображение (этап 2) в заданную библиотеку логических элементов заказной СБИС. Важнейшим является первый этап, на котором выбирается форма представления системы булевых функций и осуществляется логическая минимизация этой формы. Результат выполнения первого этапа определяет важнейшие параметры синтезированной на втором этапе логической схемы – площадь, временную задержку и энергопотребление.

I. БЛОЧНОЕ ПОКРЫТИЕ МАТРИЧНЫХ ФОРМ РАЗРЕЖЕННЫХ СИСТЕМ ДНФ БУЛЕВЫХ ФУНКЦИЙ

Рассматриваются разреженные матричные формы системы ДНФ полностью определенных булевых функций. Для таких форм систем ДНФ троичная матрица, задающая элементарные конъюнкции, содержит большую долю неопределенных значений (отсутствующих литералов булевых переменных), а булева матрица, задающая вхождения конъюнкций в ДНФ функций, содержит большую долю нулевых значений и, следовательно, небольшую долю единичных значений. В качестве базовых многоуровневых представлений использованы бинарные диаграммы решений с инверсными кофакторами (BDDI-представления) и булевы сети (Bool-представления) [1]. Минимизация BDDI-представлений выполняется по матричным заданиям систем ДНФ булевых функций,

Bool-представлений – по логическим уравнениям, задающим те же системы ДНФ.

Пусть (T, B) – пара матриц, задающая матричную форму системы ДНФ булевых функций, где T – троичная матрица, задающая общие элементарные конъюнкции, B – булева матрица, единичные элементы которой отмечают вхождения элементарных конъюнкций в ДНФ функций.

Рассмотрим пару (T_i, B_i) подматриц, где T_i – строчная (образованная некоторыми строками матрицы T) подматрица матрицы T , B_i – подматрица матрицы B , заданная на том же подмножестве строк, что и T_i . Назовем пару (T_i, B_i) блоком H_i .

Множество $(H_1, \dots, H_v) = ((T_1, B_1), \dots, (T_v, B_v))$ блоков назовем блочным покрытием пары матриц (T, B) , если каждый единичный элемент матрицы B входит только в одну из подматриц B_i , а каждая строка матрицы T входит хотя бы в одну из подматриц T_i , $i = 1, \dots, v$.

Задача блочного покрытия системы ДНФ по критерию минимальности площади блоков: найти покрытие пары (T, B) возможно меньшим числом блоков (T_i, B_i) , $i = 1, \dots, w$, имеющих возможно меньшую суммарную площадь.

Для разреженных систем ДНФ булевых функций предлагаются алгоритмы и программы [2] решения задачи блочного покрытия.

II. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Проведено экспериментальное исследование эффективности применения блочных и многоуровневых BDDI и Bool-представлений при синтезе комбинационных блоков заказных СБИС в библиотеке КМОП элементов [1]. Во всех экспериментах после логической оптимизации полученные представления систем булевых функций конвертировались в VHDL описания, синтез выполнялся в синтезаторе LeonardoSpectrum. Синтезированные схемы сравнивались по площади и временной задержке.

Эксперимент 1. Реализация исходных систем ДНФ, логическая минимизация не выполнялась.

Эксперимент 2. Для систем ДНФ функций сначала выполнялось блочное покрытие по критерию минимальности суммарной площади блоков, затем для каждого блока выполнялась совместная BDDI-минимизация.

Эксперимент 3. Для систем ДНФ функций сначала выполнялось блочное покрытие по критерию минимальности суммарной площади блоков, затем для каждого блока выполнялась совместная Bool-минимизация.

Результаты экспериментов на потоке разреженных систем ДНФ из библиотеки примеров LGSynth91 приведены в таблицах 1–3, где

Z – число литералов в задании системы булевых функций;

$Area$ – суммарная площадь элементов схемы в условных единицах;

$Delay$ – временная задержка схемы (нс);

p – ограничение на число входов блока;

r – число блоков в блочном покрытии системы ДНФ.

Символом * отмечены лучшие решения по площади и временной задержке.

III. ЗАКЛЮЧЕНИЕ

В результате проведенных экспериментов установлено: для разреженных систем ДНФ предлагаемый подход позволил в шести случаях из двенадцати улучшить решения по площади схем, получаемые синтезатором LeonardoSpectrum, и увеличить быстродействие в девяти случаях.

Таблица 1 – Эксперимент 1

Имя	Z	$Area$	$Delay$
C8	204	*21 500	*2.20
DALU	1 404	106 249	10.79
LAL	529	*26 343	3.80
PM1	124	*10 764	2.58
SCT	253	20 406	3.20
TTT2	1 263	21 500	*3.85
ALU4	7 875	487 848	9.94
APEX5	204	*188559	8.77
I2c	7 106	280 601	9.12
X1	2 148	*67 546	*3.99
X3	5 045	203 659	6.75
X4	2 649	*95 781	5.38

Таблица 2 – Эксперимент 2

Имя	Z	p	r	$Area$	$Delay$
C8	266	12	5	22 169	2.20
DALU	1 305	20	12	68 305	5.33
LAL	588	14	4	27 710	*3.25
PM1	198	8	6	11 127	*2.20
SCT	397	10	6	*19 301	*2.93
TTT2	1 106	15	7	57 189	4.64
ALU4	3 469	14	7	373 492	11.83
APEX5	6 126	40	117	195 903	*6.94
I2c	6 807	40	183	278 258	*6.56
X1	1 709	25	7	81 652	4.66
X3	4 238	25	135	229 427	6.74
X4	3 402	15	189	123 720	5.75

Таблица 3 – Эксперимент 3

Имя	Z	p	r	$Area$	$Delay$
C8	255	12	21	22314	2.20
DALU	1 030	20	34	*48044	*4.63
LAL	523	14	27	26829	3.73
PM1	168	8	20	11099	2.44
SCT	356	10	25	20116	2.95
TTT2	1 056	15	40	*41415	5.11
ALU4	3 512	14	7	*315415	11.83
APEX5	5 949	50	107	242981	6.98
I2c	4 859	40	88	*259565	7.63
X1	1 492	25	7	73087	4.71
X3	3 766	25	135	*198252	*6.67
X4	3 030	15	189	103799	*4.75

Наряду с методами минимизации систем функций в классе ДНФ эффективным методом технологически независимой оптимизации разреженных матричных форм систем ДНФ является комбинированный подход, включающий блочное покрытие системы ДНФ и последующую совместную минимизацию многоуровневых представлений блоков, при этом функции блока целесообразно минимизировать в классе булевых сетей с использованием разложений Шеннона.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Библио П.Н., Ланкевич Ю.Ю. Экспериментальное сравнение эффективности алгоритмов оптимизации BDD-представлений систем булевых функций // Программные продукты и системы. – 2020. – Т. 33. – № 3. – С. 449–463.
2. Кардаш С. Н. Построение блочных разбиений систем булевых функций на основе задачи покрытия булевых матриц // IX Международная научно-практическая конференция «BIG DATA and Advanced Analytics» (BIG DATA 2023): Материалы междунар. науч. конф., (Республика Беларусь, Минск, 17–18 мая 2023 г.). – Минск: БГУИР, 2023. – Часть 2. – С. 326 – 330.

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ПРОГРАММ МИНИМИЗАЦИИ, ВЫПОЛНЯЕМЫХ ПРИ СИНТЕЗЕ КОМБИНАЦИОННЫХ СХЕМ

Бибило П. Н., Кириенко Н. А., Романов В. И.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: {bibilo, kir, rom}@newman.bas-net.by

Приведены результаты экспериментальных исследований эффективности программ минимизации многоуровневых алгебраических представлений систем булевых функций, выполняемых при синтезе комбинационных схем. Исследуется три подхода: 1) совместная минимизация систем булевых функций; 2) раздельная минимизация; 3) выделение из исходной системы «связанных» подсистем, каждая из которых минимизируется отдельно. После минимизации выполнялся синтез логических схем в библиотеке проектирования заказных цифровых КМОП СВИС. Полученные схемы сравнивались по площади кристалла и по быстродействию. Эксперименты показали, что для некоторых систем функций эффективнее при синтезе является совместная минимизация, для других – раздельная минимизация, среднюю позицию занимает оптимизационная процедура выделения «связанных» подсистем для совместной минимизации.

ВВЕДЕНИЕ

Минимизация двухуровневых либо многоуровневых представлений систем булевых функций является первым и важнейшим этапом, выполняемым при синтезе комбинационных схем. Двухуровневыми (И-ИЛИ) представлениями называют представления функций в виде дизъюнктивных нормальных форм (ДНФ), многоуровневыми – различные формы функциональных разложений. В качестве основных методов технологически независимой оптимизации применяются методы минимизации графовых BDD-представлений [1] и их модификаций, при этом внимание чаще уделяется совместной минимизации [1, 2] и декомпозиции [3]. Одинаковость частей в таблицах истинности булевых функций привела к идее использования связанности (общности) областей определений булевых функций при синтезе многовыходных комбинационных схем. Выделение связанных подсистем функций по таблицам истинности и ДНФ функций рассматривалось в [4], где были предложены различные меры связанности функций. В данной работе предложена новая мера связанности подсистем функций и алгоритм последовательного формирования связанных подсистем как для BDD-представлений, так и для булевых сетей.

I. ВЫДЕЛЕНИЕ ПОДСИСТЕМ С ЗАДАННОЙ МЕРОЙ СВЯЗАННОСТИ

Предлагаемый эвристический алгоритм является модификацией предложенного в [4] алгоритма и состоит в последовательном формировании (на каждой итерации i) по текущей (остаточной) системе функций очередной подсистемы P_i функций (подсистема P_i характеризуется мерой компонентной связанности, не меньшей q). На первой итерации ($i=1$) текущую систему функций образуют функции исходной системы. На каждой итерации требуется выполнить шаги 1-3.

Шаг 1. Рассмотреть неупорядоченную пару функций $\{f_i, f_j\}$, $i, j = 1, \dots, m, i \neq j$, текущей системы и найти такую пару функций L , которая имеет максимальное значение меры связанности, и не меньше параметра q . Если указанной пары функций нет, то переход на шаг 4.

Шаг 2. Составить из функций найденной на первом шаге пары L подсистему P_i , и добавлять в нее поочередно те функции f_r , которые находятся с помощью следующей эвристики: из множества функций текущей системы выбирается та функция f_r , которая обеспечивает наибольшее возможное значение меры компонентной связанности для подсистемы $P_i \cup \{f_r\}$.

Шаг 3. Если нет ни одной функции f_r такой, что подсистема $P_i \cup \{f_r\}$ имеет меру компонентной связанности не меньшую q , то закончить формирование подсистемы P_i и объявить не входящие в нее функции текущей системой. Переход на шаг 1 для формирования подсистемы на итерации $i+1$.

Шаг 4. Закончить формирование подсистем, когда все функции текущей системы будут включены в формируемые подсистемы, либо когда в текущей системе нельзя будет найти ни одной пары функций, характеризующихся мерой связанности, не меньшей q , либо в текущей системе имеется только одна функция. Конец алгоритма.

Достоинством данного алгоритма является его высокое быстродействие.

II. ОПИСАНИЕ ЭКСПЕРИМЕНТОВ

В экспериментах использовалась система FLC-2 [5]. Исходными описаниями блоков комбинационной логики являлись 39 примеров систем булевых функций на языке SF [5] в виде матричных описаний систем ДНФ – SDF описаний языка SF. Исходные описания могут быть представлены булевыми сетями (Bool-представлениями) [2]. Под Bool-представлениями в системе FLC-2 понимаются булевы сети, описываемые логически-

ми уравнениями с одним логическим оператором дизъюнкции либо конъюнкции над двумя литералами булевых переменных. Были использованы следующие программы:

BDD_Builder – программа минимизации отдельных и совместных BDDI-представлений системы булевых функций;

BoolNetOpt2 – программа минимизации отдельных и совместных Bool-представлений системы булевых функций;

Splitter – программа выделения подсистем функций (при заданном значении q меры связанности) из совместных BDDI и из совместных Bool-представлений систем функций;

AutoSplit – программа нахождения лучшего значения q меры связанности по критерию минимальности общего числа литералов в BDDI- либо в Bool-представлениях систем функций.

Эксперимент 1. Совместная и отдельная минимизация в классе BDDI-представлений. При отдельной минимизации исходные матричные представления систем функций представлялись в виде логических сетей (формат BDDI). Для минимизации BDDI-представлений всей системы и отдельных функций использовалась программа BDD_Builder.

Эксперимент 2. Исходными для эксперимента 2 являлись совместно минимизированные BDDI-представления, для которых выполнялась программа Splitter для значений меры связанности $q = 10\%$, $q = 20\%$.

Эксперимент 3. Совместная и отдельная минимизация в классе булевых сетей (Bool-представлений). Базовой программой являлась программа BoolNetOpt2 [2].

Эксперимент 4. Исходными описаниями являлись совместно минимизированные Bool-представления, для которых выполнялась программа Splitter для значений меры связанности $q = 5\%$, $q = 20\%$.

Эксперимент 5. Оценка выделения связанных подсистем. Для выделения совместно минимизируемых связанных подсистем с помощью программы AutoSplit ищутся лучшие значения параметра связанности q для BDDI-представлений путем итеративного выполнения программы Splitter для меры связанности $q = 5\%$, 10% , 15% , 20% , ..., 95% , оценивается суммарное число литералов в каждой полученной логической сети и лучший вариант решения выбирается по минимуму найденной оценки.

Эксперимент 6. Совпадает с экспериментом 5, при этом исходными были совместно минимизированные Bool-представления, а результирующие

логические сети обрабатывались с помощью тех же стратегий, что и в эксперименте 3.

III. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Результаты экспериментов 1 и 2 показывают, что предложенный алгоритм выделения подсистем в 13 примерах схем (из 39) позволил получить схемы меньшей площади для исходных совместных BDDI-представлений. Получение схем с меньшей задержкой чаще было обеспечено применением отдельной BDDI минимизации и алгоритмом выделения подсистем. При выделении подсистем из совместных Bool-представлений примерно в половине случаев лучшие решения были получены алгоритмом выделения подсистем. Критерий минимальности общего числа литералов в BDDI-представлениях оказался эффективным при автоматическом выборе лучшего решения для программы Splitter. Сравнение с «ручным» выбором лучшего значения меры связанности показало погрешность в 3,9% для BDDI-представлений и 6,4% для Bool-представлений.

ЗАКЛЮЧЕНИЕ

Экспериментально установлено, что для некоторых систем функций эффективнее при синтезе логических схем является совместная минимизация, для других – отдельная минимизация, среднюю позицию занимает процедура выделения подсистем для совместной минимизации. Эффект достигается за счет того, что каждая выделенная подсистема минимизируется на основе разложения Шеннона по своей (для каждой подсистемы) перестановке переменных разложения.

1. Amaru, L. G. *New Data Structures and Algorithms for Logic Synthesis and Verification* / L. G. Amaru // Springer. – 2017. – 156 p.
2. Бибило, П. Н. Экспериментальное сравнение эффективности алгоритмов оптимизации BDD-представлений систем булевых функций / П. Н. Бибило, Ю. Ю. Ланкевич // Программные продукты и системы. – 2020. – Т. 33. – N 3. – С. 449-463.
3. Kubica, M. SMTBDD: New form of BDD for logic synthesis / M. Kubica, D. Kania // *International Journal of Electronics and Telecommunications*. – 2016. – V. 62. – № 1. – P. 33–41.
4. Бибило, П. Н. Выделение подсистем связанных функций из многоуровневого представления системы булевых функций / П. Н. Бибило, А. М. Позняк // Информатика. – 2020. – Т. 17. – N 1. – С. 63-77.
5. Бибило, П. Н. Система логической оптимизации функционально-структурных описаний цифровых устройств на основе производственно-фреймворковой модели представления знаний / П. Н. Бибило, В. И. Романов // Проблемы разработки перспективных микро- и наноэлектронных систем. – 2020. – Сб. трудов / под общ. ред. акад. РАН А. Л. Стемповского. – М. : ИПИМ РАН. – 2020. – N 4. – С. 9.

РЕГИСТРАТОР ВИБРАЦИОННЫХ СИГНАЛОВ НА ОСНОВЕ ВСТРАИВАЕМОЙ МИКРОКОНТРОЛЛЕРНОЙ СИСТЕМЫ

Бранцевич П.Ю., Лапицкая Н.В., Леванцевич В.А.
Кафедра программного обеспечения информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: branc@bsuir.edu.by, lapan@bsuir.by, lvn@bsuir.by

Анализируются технические средства контроля вибрационного состояния технических объектов. на длительном интервале времени. Обосновывается применение недорогих регистраторов вибросигналов. Предлагается архитектура микроконтроллерной встраиваемой системы для построения регистратора вибросигналов. Описываются особенности её аппаратной и программной реализации.

ВВЕДЕНИЕ

Одним из важнейших параметров, характеризующих состояние технических объектов, являются параметры вибрации, причем при длительном безостановочном функционировании технического объекта требуется непрерывное наблюдение за его вибрационным состоянием. Поэтому актуальной является задача создания приборов и систем, способных фиксировать и обрабатывать вибрационные сигналы на длительном интервале времени [1].

I. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

Простейшими приборами для оценки уровня вибрации механизмов с вращательным движением являются виброметры, измеряющие среднее квадратическое значение механических колебаний в частотном диапазоне 10 - 1000 Гц или на некоторых фиксированных частотах. Виброметры позволяют сделать общую оценку вибрационного состояния, однако они дают очень мало информации для целей диагностики.

Значительно более совершенными приборами являются спектральные анализаторы, позволяющие получить спектральный состав исследуемых вибросигналов. Такие приборы позволяют получить информацию для формулирования значительного числа диагностических признаков. Но они, в подавляющем большинстве, предназначены для проведения периодических обследований, что таит в себе опасность пропуска дефектов, проявляющих себя лишь на коротких временных интервалах с неопределенной периодичностью [2].

Одними из наиболее совершенных систем являются многоканальные стационарные измерительные комплексы. Они позволяют получать информацию различного типа при непрерывной работе и сохранять полученные данные, а также осуществлять функции предупредительной сигнализации и защиты. Такой режим работы позволяет комплексно контролировать состояние объекта и выявлять даже редко происходящие аномальные ситуации, что позволяет выявлять зарождение дефектов на самых ранних стади-

ях. Но такие системы имеют относительно высокую стоимость [2]. Поэтому актуальным является использование недорогих регистраторов вибрационных сигналов, позволяющих фиксировать вибросигналы на длительном временном интервале и сохранять их на носителе или передавать по определенному интерфейсу в мобильную или стационарную вычислительную систему для последующей обработки, анализа и принятия решений о техническом состоянии объекта.

Для этих целей удобно использовать микроконтроллерные системы на кристалле, например такие как, Atmel, Stm, ESP, и др. Данные системы построены по Гарвардской архитектуре, имеют RISC-процессор, широкий набор специализированных, периферийных устройств и низкое энергопотребление. Все это позволяет использовать микроконтроллеры для решения различных задач [3].

II. РЕАЛИЗАЦИЯ

Для разработки аппаратного модуля регистрации вибрационных сигналов к нему должны быть определены следующие основные технические требования:

1. Частотный диапазон исследуемого вибросигнала: 1 – 10000 Гц;
2. Амплитудный диапазон регистрируемых вибросигналов в единицах виброускорения: 0,1 – 1000 м/с²;
3. Количество каналов: – 1;
4. Разрядность цифроаналогового преобразователя: не менее 12 разрядов;
5. Максимальная частота дискретизации: не менее 25 кГц;
6. Относительная погрешность измерений определяется характеристиками применяемых виброизмерительных преобразователей и не должна превышать по абсолютной величине 10%;
7. Промежуточное сохранение результатов измерений на SD-карту;
8. Поддержка файловой системы FAT;

9. Поддерживаемые интерфейсы передачи данных: Serial, USB, Bluetooth, WI-FI;
10. Способы управления: автономный, внешний;
11. Питание: от внешнего бока питания и от аккумулятора.

На основе анализа технических характеристик и стоимости микроконтроллеров различных производителей было принято решение использовать микроконтроллер Stm32F103 компании STMicroelectronics, на базе процессора Cortex M3, установленного на плату прототипирования STM Nucleo board [3]. Структурная схема регистратора приведена на рисунке 1.

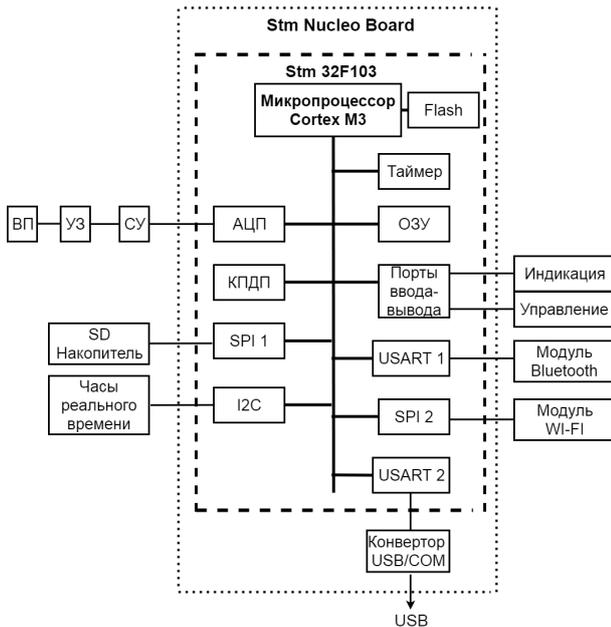


Рис. 1 – Структурная схема регистратора.

В состав регистратора входят следующие основные блоки:

- ВП – вибрационный преобразователь;
- УЗ – усилитель заряда;
- СУ – согласующий усилитель;
- АЦП – аналого-цифровой преобразователь;
- КПДП – контроллер прямого доступа к памяти;
- ОЗУ – оперативное запоминающее устройство данных;
- Flash – постоянная перепрограммируемая память команд;
- SPI, USART – контроллеры интерфейсов SPI, USART;
- Конвертор USB/COM – преобразователь интерфейсов USB/COM;
- Порты ввода/вывода – для подключения пульта управления и индикатора.

Микроконтроллер имеет встроенный 12-битный аналого-цифровой преобразователь с управляемой таймером частотой дискретизации до 1МГц. Наличие контроллера прямого доступа к памяти позволяет записывать данные с АЦП сразу в ОЗУ.

В качестве накопителя данных применяется внешний модуль SDCard, подключаемый к SPI-интерфейсу микроконтроллера. Для записи данных на SDCard используется библиотека Fatfs. Необходимый объем карты памяти можно приблизительно оценить исходя из того, что для записи 10 минут сигнала при частоте дискретизации АЦП 25кГц, требуется около 30 мегабайт памяти SDCard.

Таймер реального времени позволяет фиксировать текущее время, а также задавать временные интервалы регистрации выбросов сигналов. Управлять регистратором можно как с помощью пульта управления, так и с мобильной вычислительной системы. Для дальнейшей обработки оцифрованных сигналов они могут быть переданы на мобильную или стационарную вычислительную систему.

Регистратор обеспечивает передачу данных по интерфейсам: USB, Bluetooth в побайтном и пакетном режиме.

В пакетном режиме для определения начала и конца пакета используются специальные разделители (байт-стаффинг), а также формируется и передается контрольная сумма для проверки целостности пакета. Пакеты делятся на два типа: управляющие и пакеты данных. Управляющие пакеты содержат поле типа команды, которую надо выполнить в регистраторе.

К управляющим пакетам относятся:

- пакеты работы с файловой системой SD-карты регистратора FatFs, которые передают команды создания, открытия, записи, чтения, удаления файлов;
- пакеты работы с системным временем регистратора, передающие команды установки и получения системного времени регистратора;
- пакеты для управления устройствами регистратора.

На стороне регистратора программное средство разработано на языке C в среде Cude MX.

III. ВЫВОДЫ

Использование регистратора позволяет накапливать вибрационные сигналы на длительном промежутке времени и фиксировать изменения вибрационного состояния технического объекта.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Неразрушающий контроль. Справочник. Том 7. Книга Вибродиагностика /Ф.Я. Балицкий и др. М.: Машиностроение, 2005. – 485 с.
2. Бранцевич, П. Ю. Компьютерные системы и комплексы обработки вибрационных сигналов / П. Ю. Бранцевич. – Минск: Бестпринт, 2023. – 282 с.
3. STMicroelectronics. [Электронный ресурс]. – Электронные данные. – Режим доступа: URL:<https://www.st.com/en/evaluation-tools/stm32-nucleo-boards.html>

ПОДХОДЫ К ПОСТРОЕНИЮ ДОВЕРЕННОЙ СРЕДЫ ИСПОЛНЕНИЯ В СОВРЕМЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ ПЛАТФОРМАХ

Габер П. Н., Диваков Н. В.

Группа Микропрограммного ПО, Отдел Системного ПО, ЗАО «Инженерный Центр ЯДРО»

Минск, Республика Беларусь

E-mail: p.gaber@yadro.com, n.divakov@yadro.com

Работа посвящена вопросу построения доверенной среды исполнения (ДСИ) для купирования угроз безопасности во встраиваемых системах, построенных на наиболее распространенных современных архитектурах. В ней рассматриваются варианты реализации ДСИ на программно-аппаратном уровне, а также приводится обзор существующих решений.

ВВЕДЕНИЕ

Современные средства вычислительной техники (СВТ) для работы прикладного программного обеспечения (ПО) предоставляют расширяемую универсальную среду исполнения (УСИ). УСИ обеспечивает гибкость взаимодействия и широкие функциональные возможности, однако оставляет устройство уязвимым для широкого спектра угроз безопасности, как компрометация, изменение или утрата данных. Решением, направленным на минимизацию ущерба от реализации угроз СВТ, является применение в их составе программно-аппаратного комплекса (ПАК) с доверенной средой исполнения (ДСИ) для изоляции данных, представляющих ценность, и для выполнения операций с такими данными.

В работе рассматриваются варианты реализации ДСИ на программно-аппаратном уровне, применяемые в вычислительных платформах (ВП) на основе современных микропроцессорных архитектур.

I. УГРОЗЫ БЕЗОПАСНОСТИ ПАК

Угрозы безопасности ПАК по их цели можно разделить на угрозы, направленные на компрометацию чувствительной информации и угрозы, направленные на повреждение или уничтожение данных. По времени проведения атаки выделяют атаки времени исполнения и недокументированные возможности несанкционированного доступа (бэкдоры).

Атаки времени исполнения можно разделить на [1]:

1. Атаки на инструкции возврата/вызова (ROP/COP/JOP) используют уязвимости переполнения буфера для перезаписи инструкций возврата/вызова на стеке;
2. Атаки трансляции адреса для преодоления барьеров изоляции виртуальной памяти через подмену прав доступа используя уязвимости в модулях ядра УСИ;
3. Атаки нарушения безопасности памяти (Memory Safety Violations), например, использования памяти после освобождения, эксплуатируют известные ошибки в ПО;

4. Атаки по сторонним каналам используют анализ паттернов работы ЦПУ и не прямое воздействие на исполняемый код для поиска и извлечения данных из кэша.

II. ДОВЕРЕННАЯ СРЕДА ИСПОЛНЕНИЯ

ДСИ – это аппаратное или программно-аппаратное решение, реализующее среду исполнения, изолированную от УСИ, но имеющую канал для взаимодействия с УСИ.

ДСИ должна обеспечивать:

1. Безопасность активов, управляемых ДСИ;
2. Безопасное размещение и исполнение ДП, их изоляцию друг от друга;
3. Безопасное хранение конфиденциальных данных с обеспечением согласованности, целостности и привязки к ДСИ;
4. Защищенный коммуникационный канал связи между КП в УСИ и ДП в ДСИ, включая конечные точки в ДСИ;
5. Сервисы контроля целостности кода и данных УСИ и ее приложений.

В качестве безопасного контура ДСИ должна обеспечивать выполнение вычислений, критических для безопасного функционирования устройства и УСИ, а также конфиденциальность и целостность обрабатываемых данных.

ДСИ должна нейтрализовать следующие угрозы СВТ в случае компрометации УСИ:

1. Несанкционированный доступ и изменение критически важной информации, вмешательство в работу компонентов УСИ;
2. Запуск на исполнение недоверенных приложений;
3. Соккрытие попыток и фактов компрометации системы.

III. ВАРИАНТЫ ПОСТРОЕНИЯ ДСИ НА ВП

ДСИ может быть реализована различными способами в зависимости от архитектуры, области применения, и сложности СнК.

Вариант 1. ДСИ размещается на отдельной СнК со своими собственными аппаратными ресурсами (память, контроллеры ввода-вывода и т.п.) (см. рис. 1). Такой вариант реализации ДСИ

обеспечивает максимальный уровень изоляции компонентов ДСИ от УСИ и максимальный уровень безопасности.

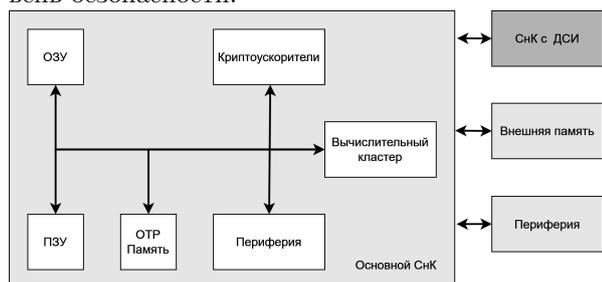


Рис. 1 – Размещение ДСИ на отдельной СнК

Вариант 2. ДСИ с использованием общих ресурсов с УСИ (см. рис. 2). Такой вариант предполагает что ДСИ выполняется на той же ВП, что и УСИ и совместно использует его аппаратные ресурсы. Разделение ДСИ и УСИ выполняется механизмами уровней безопасности (запуска), различными для ДСИ и приложений УСИ.

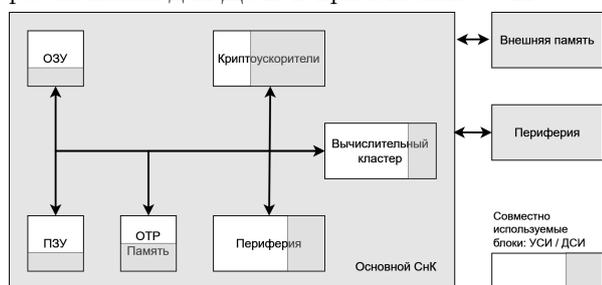


Рис. 2 – Совместное использование ДСИ и УСИ ресурсов СнК

Вариант 3. ДСИ с выделенным вычислительным ядром и аппаратными ресурсами (см. рис. 3). В этом случае для ДСИ выделяется одно из ядер СнК со своим набором ресурсов в изолированной подсистеме, которая имеет доступ к аппаратным ресурсам УСИ.

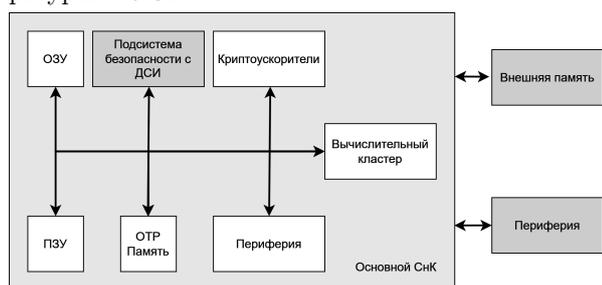


Рис. 3 – Выделенное ДСИ на общем СнК с УСИ

IV. РЕАЛИЗАЦИИ ДСИ В СОВРЕМЕННЫХ АРХИТЕКТУРАХ ВЫЧИСЛИТЕЛЬНЫХ СРЕДСТВ

X86. В решениях компаний AMD и Intel используются различные подходы к реализации ДСИ. Intel Software Guard Extension (SGX) – это набор инструкций, который могут использоваться приложениями для создания защищенных областей кода и данных [2]. AMD Platform

Security Processor (PSP) представляет собой сопроцессор на базе ARM (Cortex A5) с расширением TrustZone. Сопроцессор начинает работу до запуска основного вычислительного кластера и отвечает за верификацию UEFI загрузчика, его запуск и предоставление TPM сервисов [3].

ARM. Расширения безопасности ARM TrustZone позволяют разделять ресурсы и вычислительные ядра СнК на аппаратном уровне и создавать на их основе ДСИ. Изоляция КП и ДП выполняется с использованием защищенной памяти, контроллера прерываний, механизмами изоляции памяти и уровней безопасности (от EL0 до EL3) приложений [4].

RISC-V. На данный момент архитектура RISC-V не имеет стандартного механизма реализации ДСИ. На рынке присутствует несколько распространенных решений на базе RISC-V, которые базируются на расширении Physical Memory Protection (RISC-V PMP), которое использует специальные регистры для обеспечения механизма изоляции страниц памяти. SiFive WorldGuard решение является альтернативой ARM TrustZone. Оно использует специальные регистры для контроля прерываний, супервизор безопасности и три уровня исполнения кода для изоляции ДП от КП [5].

MIPS. MIPS TEE использует расширение виртуализации (VZ) и гипервизор L4Re для изоляции ДСИ (TEE) от УСИ (REE).

ЗАКЛЮЧЕНИЕ

В работе рассмотрены требования к реализации ДСИ, типовые варианты реализации и наиболее распространенные решения ДСИ на базе современных микропроцессорных архитектур.

1. Hardware-Enabled Security [Electronic Resource] / M. Bartock, M. Souppaya [and others]. – National Institute of Standards and Technology, 2022. – Mode of access: <https://nvlpubs.nist.gov/nistpubs/ir/2022/Nist.IR.8320.pdf>. – Date of access: 17.03.2023.
2. The Intel SGX Memory Encryption Engine [Electronic Resource] / S. Jonson. – Intel Software., 2016. – Mode of access: <https://software.intel.com/en-us/blogs/2016/02/26/memory-encryption-an-intel-sgx-underpinning-technology>. – Date of access: 17.10.2023.
3. AMD Platform Security Processor [Electronic Resource] / Wikipedia. – Mode of access: https://en.wikipedia.org/wiki/AMD_Platform_Security_Processor. – Date of access: 17.10.2023
4. Самоделов, А. Аппаратная поддержка доверенной среды исполнения в микроконтроллерах и микропроцессорах с архитектурой ARMv.8A и ARMv.8M / А. Самоделов // Материалы XXII научно-практической конференции «РусКрипто'2020». – 2020.
5. SiFive WorldGuard Technical Paper [Electronic Resource] / SiFive, 2023. – Mode of access: https://sifive.cdn.prismic.io/sifive/f5dcaa9d-a0fd-4d91-b5e6-9ad4e5930c5e_WorldGuard-Technical-Paper_v2.4.pdf. – Date of access: 17.10.2023.

АНАЛИЗ ПРИМЕНЕНИЯ ПРОГРАММНЫХ ЭМУЛЯТОРОВ В МАРШРУТЕ ПРОЕКТИРОВАНИЯ СнК

Кишко Д. В., Алехин А. П.
ООО Ядро Микропроцессоры
Москва, Россия
E-mail: d.kishko@yadro.com

В работе проведен сравнительный анализ программных эмуляторов для систем на кристалле (СнК). Рассмотрены способы интеграции эмуляторов с основными инструментами разработки в маршруте проектирования и верификации СнК. Показано, что представленный подход позволяет снизить время на разработку и отладку функциональных сценариев за счет включения в работу на ранних стадиях проектирования и переноса наиболее затратных процедур отладки в эмулятор.

ВВЕДЕНИЕ

Для обеспечения высокого качества проектирования СнК требуются значительные усилия в области функциональной верификации дизайна интегральной схемы. Процесс верификации позволяет определить соответствие разрабатываемого дизайна предъявляемым требованиям, и выявить на ранних этапах скрытые проблемы функционирования как отдельных частей СнК, так и в целом интегральной схемы. Исследования последних лет показывают, что более 50% времени дизайн-компании затрачивают на процессы верификации СнК [1]. Существуют разные подходы к функциональной верификации СнК, отличающиеся сложностью, стоимостью и глубиной осуществляемых проверок. Согласно исследованиям, представленным в [2] и [3], к методам функциональной верификации относят: формальную верификацию, поведенческую, потактовую и событийную симуляцию (HDL симуляцию), использование аппаратных ускорителей и прототипирования на ПЛИС. Сам процесс верификации включает не только планирование, разработку и отладку функциональных сценариев, но также запуск этих сценариев на выполнение, который может занимать значительное время.

Программные эмуляторы СнК позволяют значительно снизить общее время отладки, запуска и выполнения функциональных сценариев, в связи с этим представляют значительный интерес для внедрения в маршрут проектирования СнК.

I. МАРШРУТ ПРОЕКТИРОВАНИЯ СнК

Обобщенный маршрут проектирования СнК представлен на рисунке 1. Как правило, он включает в себя следующие этапы: исследование рынка и формирование продуктовых требований к СнК, архитектурную проработку, этапы логического и физического дизайна, выпуск СнК, процедуры отбраковки и приемки изготовленной интегральной схемы. Основные задачи функциональной верификации СнК начинаются на этапе проработки архитектуры и заканчиваются началом этапа проработки физического дизайна.



Рис. 1 – Обобщенный маршрут проектирования СнК

Как правило разработка и верификация самой программной модели СнК происходит значительно быстрее, чем разработка логического дизайна СнК. При этом использование программных эмуляторов позволяет сократить время на функциональную верификацию СнК ввиду того, что отладка функциональных сценариев на эмуляторе занимает гораздо меньше времени, чем отладка с использованием HDL симулятора. Это достигается в первую очередь за счет применения высокоуровневых языков программирования, использования готовых библиотечных моделей составных частей СнК. Также большинство программных эмуляторов системного уровня позволяют осуществить запуск целевой операционной системы и произвести валидацию пользовательских сценариев, что является дополнительным преимуществом в повышении качества проектируемого дизайна. На рисунке 2 представлена взаимосвязь процессов разработки логического дизайна, функциональных сценариев и программной модели.

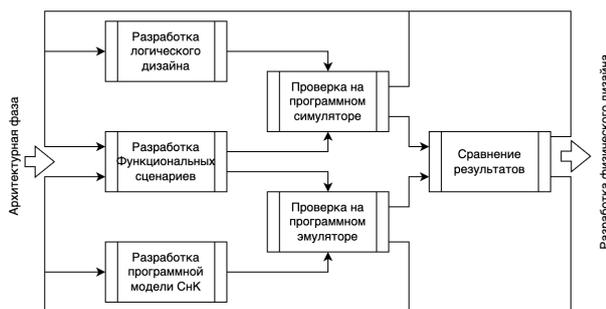


Рис. 2 – Процессы разработки и функциональной верификации СнК

Создание программной модели разрабатываемого СнК, позволяет решить ряд задач, основными из которых являются:

- Разработать функциональные сценарии использования СнК, необходимые для верификации логического дизайна;
- Верифицировать функциональные сценарии;
- Выработать требования к программной модели функциональных блоков СнК;
- Сформировать артефакты (отладочный вывод, трассы выполнения CPU) для сравнения с результатами работы симулятора.

II. АНАЛИЗ СОВРЕМЕННЫХ ЭМУЛЯТОРОВ

В начале разработки логического дизайна СнК его программная модель может иметь достаточно ограниченный функционал, позволяющий проводить минимальные функциональные проверки. Однако в процессе проектирования СнК и доработки его модели, перечень проверок расширяется, обеспечивая тем самым полноту проводимых проверок. Это позволяет подойти к концу этапа логического дизайна с практически полным функционально верифицированным дизайном СнК. Неоспоримым преимуществом эмуляторов является возможность их совместного использования вместе с HDL симуляторами с помощью DPI, а также интеграция с аппаратными ускорителями на базе FPGA [3].

В зависимости от сложности проектируемого СнК к эмуляторам могут предъявляться различные требования, основные из которых представлены ниже:

- T1. Открытый исходный код;
- T2. Быстродействие эмулятора;
- T3. Возможность создания модели вычислительного кластера, состоящего из большого количества ядер, в том числе гетерогенных;
- T4. Возможность создания и подключение моделей периферийных устройств;
- T5. Быстрая и гибкая настройка адресного пространства и карты прерываний;
- T6. Полная документация и наличие методических материалов;

- T7. Гибкость получения трасс выполнения CPU и отладочной информации;
- T8. Возможность сохранения среза архитектурного состояния СнК;
- T9. Возможность программной отладки (GDB).

В таблице 1 представлены сравнительные характеристики программных эмуляторов СнК. Знаком + показано соответствие требованию, знаком – несоответствие требованию, и частичное соответствие – знаком ±.

Таблица 1 – Сравнение платформ эмуляции СнК

Платформа	T1	T2	T3	T4	T5	T6	T7	T8	T9
Simics	–	+	+	+	+	+	±	+	+
OVPSim	±	+	+	+	+	+	±	+	+
QEMU	+	+	–	±	±	–	±	–	+
Spike	+	±	+	+	+	–	+	–	+
riscv-rust	+	±	±	+	+	+	+	–	–
tinyemu	+	±	+	–	–	+	+	–	–
gem5	+	+	+	–	–	–	–	–	–
whisper	+	–	+	–	–	–	+	–	–
RVVM	+	±	–	±	±	–	±	–	–
nemu	+	–	+	–	–	–	+	–	–
Renode	+	+	+	+	+	+	+	+	+

Разработка программной модели СнК позволяет не только внести существенный вклад в верификацию СнК и проверить различные функциональные сценарии, но и на ранних этапах проработать программную реализацию драйверов для внутренних узлов будущей интегральной схемы и при ее выпуске обеспечить готовность программных библиотек и инструментов для заданных условий эксплуатации.

III. ЗАКЛЮЧЕНИЕ

В работе представлен анализ использования программных эмуляторов применительно к маршруту проектирования СнК, рассмотрены особенности применения в совокупности с предъявляемыми требованиями. В зависимости от подходов применяемых для верификации анализ позволяет на начальной стадии выбрать наиболее удобный инструмент для создания программной модели СнК.

IV. СПИСОК ЛИТЕРАТУРЫ

1. Wilson Research Group functional verification study IC/ASIC functional verification trend report 2020. Mode of access: <http://www.inform.ind.edu/PBIO/>. – Date of access: 10.10.2023.
2. Слинкин Д. И. Анализ современных методов тестирования и верификации проектов сверхбольших интегральных схем. Международный журнал Программные продукты и системы. – 2000. – No 10. Режим доступа: – Режим доступа: <http://www.swsys.ru/index.php?page=article&id=4308&lang=> – Дата доступа: 10.10.2023.
3. Солодовников А. П., Переверзев А. Л., Силантьев А. М. Программно-аппаратный комплекс для ускорения функциональной верификации систем на кристалле // Изв. вузов. Электроника. 2023. Т. 28. No 4. С. 441-451. <https://doi.org/10.24151/1561-5405-2023-28-4-441-451>. – EDN: NULBAQ

ИНСТРУМЕНТ ДЛЯ ОЦЕНКИ КАЧЕСТВА ПОКРЫТИЯ ИНСТРУКЦИЙ КОМПИЛЯТОРАМИ ДЛЯ АРХИТЕКТУРЫ RISC-V

Лекомцев А. А., Сухарев А. Д.
Кафедра системного программирования
Санкт-Петербургский государственный университет
Санкт-Петербург, Российская Федерация
E-mail: alelekomtsev@gmail.com, therain.i@yahoo.com

В работе рассматривается задача анализа кодогенерации компиляторов и инструмент, разработанный для автоматизации ее решения

ВВЕДЕНИЕ

RISC-V – открытая расширяемая ISA (Instruction Set Architecture) с модульной структурой. Архитектура является результатом общемирового сотрудничества большого количества компаний, входящих в RISC-V International [1]. Спецификации доступны для свободного и бесплатного использования, что способствует формированию активного сообщества. Отличительной чертой архитектуры является ее простота в сравнении с другими ISA. Стандартный набор команд содержит только 47 базовых инструкций [2]. Добавление функциональности достигается за счет разработки расширений, часть которых входит в стандарт. Благодаря такой модульности обеспечивается высокая расширяемость архитектуры. Производителям аппаратного обеспечения требуется поддерживать только базовый набор, реализация дополнительных инструкций является опциональной.

Примерами дополнительных операций являются умножение и деление целых чисел, атомарные инструкции, векторные операции, операции с числами с плавающей запятой, инструкции для работы с байтами и отдельными битами. Например, bitmanip [3] – стандартный набор расширений RISC-V, который включает в себя группы инструкций, объединенные по их назначению: инструкции для быстрого вычисления адресов, для операций с битами и другие.

Расширение дает возможность использовать вместо нескольких базовых инструкций одну, что зачастую более оптимально как по времени исполнения, так и по энергопотреблению [3]. В связи с тем, что некоторые расширения были введены в стандарт относительно недавно, компиляторы не в полной мере поддерживают распознавание добавленных инструкций в исходном коде.

Поддержка компиляторами инструкций из многочисленных расширений и их использование для оптимизации является актуальной задачей в сообществе.

I. МОТИВАЦИЯ

Одна из часто встречающихся задач в программировании – задача конвертации чисел из

представления LE в BE, и наоборот. В стандартном расширении RISC-V bitmanip присутствует инструкция rev8, которая инвертирует порядок байтов в переданном ей регистре [3]. Напишем код, решающий задачу, на языке C (Листинг 1) и сравним способность различных компиляторов транслировать его в rev8.

Сравнение результатов кодогенерации Clang 17.0.1 (Листинг 2) и GCC 13.2.0 (Листинг 3) с параметром `-march=rv64gc_zbb` и уровнем оптимизации `O2` показывает, что Clang транслирует код в rev8, когда как GCC не генерирует данную инструкцию.

```
#include <stdint.h>

uint64_t rev8(uint64_t rs) {
    uint64_t out = 0;
    for (int i = 0; i < 8; i++) {
        uint8_t tmp = rs >> (i * 8);
        out |= (uint64_t) tmp << ((7 - i) * 8);
    }
    return out;
}
```

Листинг 1 – Реализация алгоритма инверсии порядка байтов в 64 битном числе

```
rev8:
    rev8    a0, a0
    ret
```

Листинг 2 – Код, сгенерированный Clang 17.0.1

```
rev8:
    mv     a2, a0
    li     a4, 0
    li     a0, 0
    li     a6, 56
    li     a1, 64
.L2:
    srl   a5, a2, a4
    subw a3, a6, a4
    andi a5, a5, 255
    sll  a5, a5, a3
    addiw a4, a4, 8
    or   a0, a0, a5
    bne a4, a1, .L2
    ret
```

Листинг 3 – Код, сгенерированный gcc 13.2.0

Возникает потребность в сравнении возможностей кодогенерации различных компиляторов,

а также отслеживании соответствующих изменений.

Нами не было найдено инструментов, автоматизирующих этот процесс и подходящих под данные требования. Ближайшими аналогами можно назвать наборы тестов и утилиты для их запуска, используемые в компиляторе GCC и проекте LLVM. Основная проблема, с которой мы столкнулись при использовании данных инструментов – разный формат тестов. Например, тестируемый код в проекте LLVM написан на языке LLVM IR [4], тогда как в GCC – на языке C [5], что усложняет сравнение. Следовательно, было принято решение разработать инструмент для анализа возможностей кодогенерации различных компиляторов с единым форматом тестов.

II. РАЗРАБОТКА ИНСТРУМЕНТА

Основными требованиями для инструмента стали поддержка единого формата тестов, а также возможность сравнения кодогенерации различных компиляторов на разных уровнях оптимизации. В качестве дополнительных требований были выдвинуты:

- гибкость конфигурации;
- скорость выполнения;
- простота добавления функциональности;
- кросс-платформенность.

Для тестируемого кода был выбран язык C, так как за основу тестируемых компиляторов были взяты Clang и GCC. Также было принято решение структурировать тесты по каталогам, соответствующим отдельным инструкциям, что позволило писать несколько тестов для одной инструкции. Для удобного сравнения возможностей кодогенерации компиляторов формат CSV был выбран в качестве основного представления результатов тестирования.

Для написания инструмента был выбран язык программирования Python, что обеспечило высокую скорость разработки и кроссплатформенность проекта.

Высокая скорость выполнения тестов была достигнута за счет реализации асинхронного исполнения с использованием библиотеки `asyncio` из стандартной библиотеки Python. В сравнении с синхронным запуском тестов такой подход обеспечил трехкратный прирост в производительности инструмента.

Для конфигурации инструмента был выбран формат json. В конфигурационном файле пользователь может установить каталог, в котором находятся тесты, путь до тестируемого компилятора, а также параметры его запуска и набор интересных уровней оптимизации. Так достигается высокая гибкость конфигурации тестирования.

III. ПРИМЕР РАБОТЫ ИНСТРУМЕНТА

Для проверки возможностей кодогенерации необходимо было написать ряд тестов. В качестве тестовых случаев было принято решение использовать инструкции из расширения `bitmanip`. В документации для каждой инструкции этого расширения есть псевдокод, который описывает семантику инструкции [3]. Этот псевдокод был адаптирован для отдельных инструкций, а также были описаны дополнительные случаи, которые компилятор может транслировать в инструкции расширения. Тестируемыми компиляторами были выбраны Clang 17.0.4 и GCC 13.01.0, учитывались все уровни оптимизации. Пример результатов тестирования был добавлен в репозиторий в каталог `example` [5].

Для тестирования было составлено 96 тестовых сценариев, которые задавались одним из 24 шаблонов и одним из 4 уровней оптимизации компилятора. Clang 17.0.4 породил ожидаемую инструкцию в 45 случаях, GCC 13.01.0 – в 39 случаях.

ЗАКЛЮЧЕНИЕ

Некоторые расширения RISC-V были введены в стандарт относительно недавно, а соответственно компиляторы не в полной мере поддерживают распознавание новых инструкций в исходном коде программ. В сообществе актуальна задача сравнения возможностей кодогенерации различных компиляторов и их версий для инструкций из различных расширений.

Разработанный инструмент помогает автоматизировать процесс такого сравнения и анализа, предоставляя единый формат тестов с возможностью гибкой конфигурации.

СПИСОК ЛИТЕРАТУРЫ

1. RISC-V International Members <https://riscv.org/members/> [Электронный ресурс] / URL: <https://riscv.org/members/> – 2023
2. The RISC-V Instruction Set Manual [Электронный ресурс] / URL: <https://riscv.org/wp-content/uploads/2017/05/riscv-spec-v2.2.pdf> – 2023
3. RISC-V Bit-manipulation A, B, C and S Extensions [Электронный ресурс] / URL: <https://five-embeddev.com/riscv-bitmanip/draft/bitmanip.html> – 2023
4. Электронный ресурс / URL: <https://github.com/llvm/llvm-project/blob/main/llvm/test/CodeGen/RISCV/rv32zba.ll> – 2023
5. Электронный ресурс / URL: <https://github.com/GCC-mirror/gcc/blob/master/gcc/testsuite/gcc.target/riscv/zba-shadd.c> – 2023
6. Репозиторий с реализацией инструмента [Электронный ресурс] / URL: <https://github.com/vacmannnn/riscv-check> – 2023

ДИАГРАММА ГАНТА КАК СРЕДСТВО ВИЗУАЛИЗАЦИИ РЕЗУЛЬТАТОВ ПАРАЛЛЕЛЬНЫХ ПРОГРАММ

Логинова И. П.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: irilog@mail.ru

В работе представлен подход, предназначенный для применения таких инструментов визуализации как диаграммы Ганта при анализе производительности параллельных решений различных задач логической оптимизации. Подход может применяться для поиска характеристик исходных объектов, которые приводят к получению эффективного решения логико-комбинаторных задач оптимизации.

ВВЕДЕНИЕ

Решение некоторых задач проектирования СБИС сверхбольших размерностей невозможно без использования масштабируемых параллельных алгоритмов, ориентированных на многопроцессорные/многоядерные вычислительные системы. Большая часть алгоритмов из области логического проектирования имеют свои программные реализации, которые используются в составе различных САПР. Как правило, эти программы, представлены последовательно исполняемым кодом без привязки к архитектуре используемых вычислительных средств. В последние годы были приложены определенные усилия по модернизации существующих алгоритмов задач логического проектирования и разработке новых, параллельных реализаций этих алгоритмов. Результатом проведенных исследований стало признание того, что большинство логико-комбинаторных алгоритмов не допускают эффективного распараллеливания.

I. ПАРАЛЛЕЛЬНЫЙ ЗАПУСК ПРОГРАММ

Один из вариантов организации распараллеливания для программ проектирования (включенных в состав некоторых САПР) для многоядерной системы с общей памятью на основе стандарта OpenMP предложен в [1]. OpenMP является открытым стандартом для реализации многопоточных приложений, содержит набор различных директив компилятора, функций соответствующей библиотеки и переменных окружения. Стандарт OpenMP можно рассматривать как высокоуровневую надстройку над POSIX или Windows Threads (библиотеками потоков). При проектировании параллельных программ требуется выполнение следующих действий: разделение (декомпозиция) объекта проектирования на части, выявление информационных зависимостей между частями и распределение работы с этими частями по вычислительным элементам (в системах с общей памятью - по ядрам процессора). В основу подхода [1] положен запуск исполняемых модулей программ и организация их параллельной работы. Поскольку любая программа, загруженная в память компьютера для выполнения, образу-

ет процесс, то программа, реализующая запуск нескольких процессов, будет работать быстрее, если организовать параллельное выполнение запущенных процессов. Это может быть осуществлено путем вставки директив OpenMP в код программы, запускающей в цикле выполнение нескольких процессов. Такую параллельную работу с программами можно реализовать, если: 1) выполняется множество независимых процессов; 2) процессам не надо общаться; 3) результаты, полученные процессами, сводятся в один когда завершат работу все процессы.

II. АРХИТЕКТУРА ПРОГРАММЫ ДЛЯ ОРГАНИЗАЦИИ ПАРАЛЛЕЛЬНОЙ РАБОТЫ

Схема реализации параллельных вычислений такова: в головной программе запускается множество параллельных задач (процессов); эти задачи динамическим образом распределяются по всем ядрам (включая логические процессоры, если они есть); каждая задача определяет работу одного процесса. У всех процессов свои адресные пространства, процессы взаимодействуют с памятью независимо, каждый процесс работает со своими входными данными. Когда головная программа производит запуск другой программы, создается новый (дочерний) процесс. Работа дочернего процесса может осуществляться в синхронном и асинхронном режиме. Размещение директивы `pragma omp parallel for` перед циклом, который запускает группу дочерних процессов, распределяет эти процессы по параллельным регионам — каждый процесс в отдельный параллельный регион. Все потоки в параллельных регионах выполняют, как правило, один код. Этот код представлен командой `start`, которая производит запуск консольной реализации какой-либо внешней программы [1]. Аргументами команды `start` являются параметры командной строки запускаемой программы. В конце параллельной области происходит неявная барьерная синхронизация, т.е. все потоки останавливаются до тех пор, пока последний поток не выполнит свою программу целиком. Но при вставке в цикл директивы `pragma omp task` синхронизация потоков отменяется, создается набор задач (`task`), число задач определяется числом процессов, все задачи помещаются в оче-

редь. Из очереди задачу (процесс) может взять один из рабочих потоков для выполнения. Задача может быть намного больше числа рабочих потоков. Распределение задач по потокам осуществляется планировщиком задач, принцип работы которого представлен на рисунке 1. Планировщик реализует формирование очереди, проводит распределение задач по рабочим потокам, применяя стратегию work-stealing. Суть этой стратегии заключается в том, что освободившийся рабочий поток при отсутствии ожидающих задач в своей очереди заимствует задачу у одного из рабочих потоков.

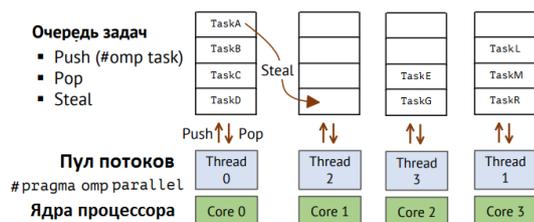


Рис. 1 – Распределение очереди в пуле потоков

III. ОРГАНИЗАЦИЯ ВИЗУАЛИЗАЦИИ

Для оценки производительности параллельных вычислений следует провести оценку времени выполнения каждого процесса - получить времена начала и завершения дочерних процессов. Поэтому в тело цикла запуска процессов вводятся так называемые критические секции (pragma omp critical), которые содержат код для вычисления и сохранения в трассировочном файле характеристик каждого процесса в порядке наступлений событий в процессе и рабочем потоке. Псевдокод организации работы параллельных процессов приведен на рисунке 2.

```
#pragma omp parallel for
// цикл по числу NX запусков внешней программы
for(i=0; i<NX; i++) {
# pragma omp task .....
{
// вычисление времени начала задачи
# pragma omp critical
{ ... }
start(process); // запуск внешней программы
// вычисление времени завершения задачи
# pragma omp critical
{ ... }
} }
}
```

Рис. 2 – Псевдокод параллельного пуска процессов

После завершения вычислений проводится упорядочение (с использованием алгоритмов рекурсивной сортировки) данных трассировочного файла. В целом, даже в упорядоченном виде данные из трассировочного файла - время, идентификаторы задачи, потока, имя программы, достаточно тяжелы для усваивания этой информации, поэтому удобным инструментом при анализе ре-

зультатов работы параллельной программы является визуализация упорядоченных данных в виде диаграмм Ганта. Обычно, как инструмент визуализации, диаграмма Ганта используется при планировании проектов и представляет из себя столбчатую диаграмму с накоплением, которая показывает даты начала, окончания, сроки проекта. Использование этого инструмента в контексте организации параллельных вычислений дает визуальное представление о времени выполнении задач. На рисунке 3 приведен вид диаграммы Ганта, полученной в результате следующего эксперимента. Было проведено параллельное выполнение трех разных программ логической оптимизации, каждая программа работала со своим объектом проектирования, каждый объект, был разбит на группы независимых частей с использованием методов декомпозиции - в общей сложности, параллельно выполнялись 84 задачи оптимизации.

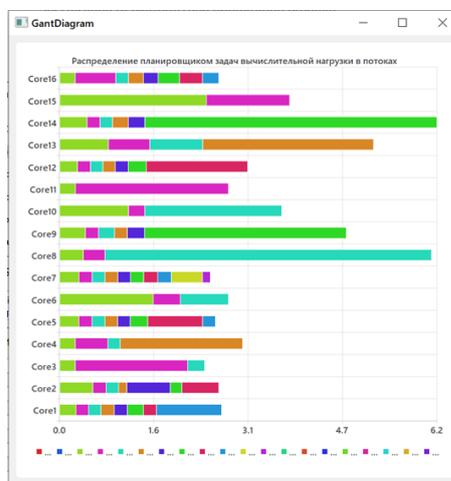


Рис. 3 – Визуализация результатов программы

Данная диаграмма отображает распределение по 16-ти процессорам этих задач. По диаграмме видно, что время выполнения отдельных задач существенно различается, что обусловлено разными алгоритмами оптимизации и разными размерами исходного объекта для каждой задачи. В итоге, анализ упорядоченных данных трассировочного файла совместно с визуальным представлением этих данных в виде диаграмм позволяет оценить влияние различных методов декомпозиции объектов проектирования, а также сравнить быстродействие программ оптимизации. Таким образом, диаграмма Ганта представляется удобным средством визуально-структурированного представления данных при анализе результатов параллельной работы программ проектирования. Программа визуализации диаграмм Ганта реализована при использовании кроссплатформенной библиотеки QtCharts (включена в Qt5.7).

1. Бибило, П. Н. Экспериментальное сравнение эффективности программ минимизации систем булевых функций в классе дизъюнктивных нормальных форм / П. Н. Бибило, И. П. Логинова // Информатика. – 2022. – Т. 19, № 2.– С. 26-55.

ГЕНЕРАЦИЯ СЛУЧАЙНЫХ ЧИСЕЛ НА ОСНОВЕ ДРЕБЕЗГА КОНТАКТОВ

Можейко Д. О., Иванюк А. А.

Кафедра электронных вычислительных машин, кафедра информатики,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: dimamozheiko13@gmail.com, ivaniuk@bsuir.by

В работе рассматривается вопрос генерации случайных чисел на основе дребезга контактов и их нормализации для достижения распределения близкого к равномерному. Предлагается схемотехническое решение для построения генератора случайных последовательностей на основе дребезга контактов. Прототипирование разработанного генератора проводилось на отладочных платах Digilent ZYBO Z7-10, цифровые схемы проектировались на языке VHDL для кристалла FPGA Xilinx Zynq-7000.

ВВЕДЕНИЕ

Дребезг контактов электромеханических коммутационных устройств неизбежное явление, наблюдаемое во время замыкания и размыкания контактов, что доказывалось в предыдущих исследованиях [1]. Причиной данного явления является кратковременное соударение и неконтролируемые отскоки контактов друг от друга, что приводит к нежелательным замыканиям и размыканиям электрических цепей.

Функционально движковые переключатели являются логическими переключателями, в то время как функционал нажатия кнопок более разнообразный: нажатие и удержание в течении заданного промежутка времени (инициализация/сброс), однократное нажатие и отпускание (настройка режима/навигация), многократное последовательное нажатие определенное число раз и различные другие комбинации.

Само явление дребезга носит случайный характер, обусловленный большим числом факторов: свойства материалов, из которых изготовлены коммутационные устройства и их контакты; геометрические размеры контактов и зазоров между ними; значения напряжения и силы тока в коммутируемых линиях; свойства окружающей среды (температура, влажность и т.п.), сила и скорость нажатия/переключения; время удержания и т.д. [2,3].

В работе рассматривается применение дребезга контактов кнопок, клавишных и движковых переключателей для генерации действительно случайных последовательностей.

1. ЦИФРОВАЯ СХЕМА ГЕНЕРАТОРА СЛУЧАЙНЫХ ЧИСЕЛ

В ходе работы был спроектирован аппаратно-программный комплекс, при этом управление генератором и передачей данных от него осуществляется встроенным процессором ARM Cortex-A9, для которого была написана соответствующая программа. Весь аппаратно-программный комплекс был спроектирован при

помощи САПР Xilinx Vivado/Vitis с применением языков VHDL и C.

Общая структура разработанного анализатора приведена на рисунке 1. Анализируемый сигнал от коммутатора подключается ко входу BTN и далее сэмплируется на триггере DFF на частоте FCLK=125 МГц [4]. Устройство управления CTRL_FSM определив первый значимый фронт сигнала от коммутатора запускает счетчик временного окна измерения TMW_CNT. Значение самого окна задается 32-разрядным значением на входной шине TMW. В этом окне измерений, в зависимости от логического состояния сигнала на линии BTN, работают линейные сдвиговые регистры с обратной связью LFSR. На LFSR_N, тактирующийся FCLK, поступает состояние сигнала BTN, учитывающееся при обратной связи, а на LFSR_B поступает логическое состояние BTN уже в качестве тактирующего сигнала. По окончании измерения блок CTRL_FSM прекращает управление счетом и вырабатывает сигнал готовности READY. После этого значения с двух LFSR складываются и передаются для дальнейшего анализа.

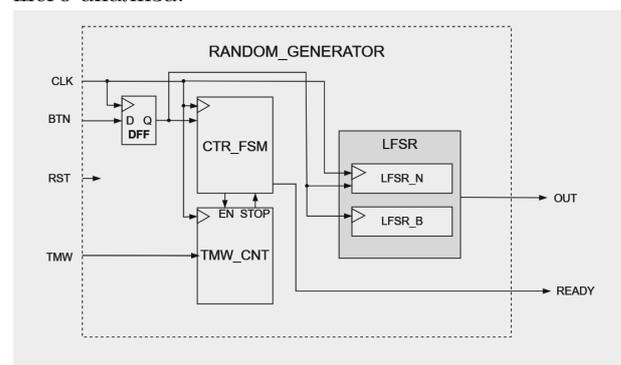


Рис. 1 – Структурная схема генератора

Использование представленной схемы, основанной на двух LFSR, вместо одного, обусловлено тем, что LFSR_B, тактирующийся сигналом BTN, будет срабатывать при каждом его фронте, тем самым реагировать на дребезг контактов коммутатора. LFSR_N, тактирующийся FCLK, будет сэмплировать сигнал BTN с частотой 125

Мгц и реагировать во время активного состояния коммутатора, такого как дребезг или его удержание в замкнутом состоянии. В итоге мы получаем схему с двумя LFSR, которая использует два фактора случайности, таких как количество фронтов сигнала и длительность активного состояния коммутатора.

II. АВТОМАТИЗАЦИЯ И РЕЗУЛЬТАТ ЭКСПЕРИМЕНТА

Для получения достаточной выборки было принято решение об автоматизации эксперимента. Была установлена система умного дома HAS (Home assistant service), налажена схема взаимодействия между HAS и устройством Fingerbot с помощью Zigbee радиомодуля. В ходе работы были рассмотрены различные типы коммутаторов и в качестве примеров выбраны: имеющаяся на плате кнопка ZYBO-Z7 (кнопка 1), а также внешний коммутатор DS-314 (кнопка 2), представленные на рисунке 2.



Рис. 2 – Рассмотренные в эксперименте коммутаторы и устройство Fingerbot

Было запрограммировано устройство Fingerbot с целью автоматизированного нажатия на кнопки в количестве 2500 раз. На рисунке 3 и 4 можно увидеть распределение величин в диапазоне от 0 до 255 для кнопки 1 и кнопки 2 соответственно.

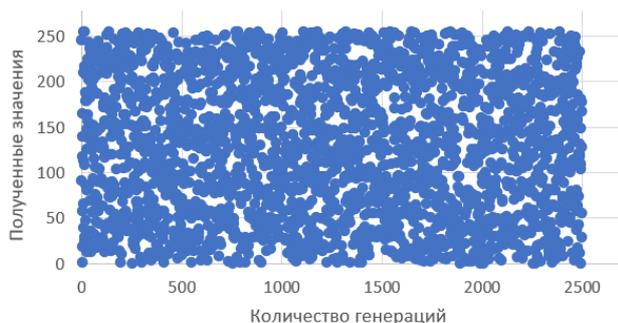


Рис. 3 – Распределение полученных значений для кнопки 1

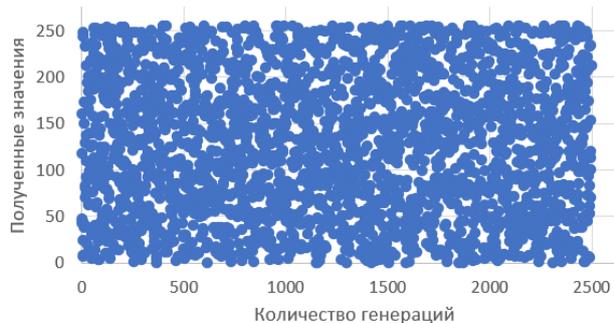


Рис. 4 – Распределение полученных значений для кнопки 2

III. ЗАКЛЮЧЕНИЕ

В работе рассмотрена генерация случайных чисел на основе дребезга контактов коммутаторов, используемых в цифровых устройствах различного назначения. В ходе работы была спроектирована цифровая система генератора случайных чисел на основе FPGA, позволяющая генерировать случайные последовательности по результатам нажатия/переключения коммутатора, учитывающая время дребезга, во время включения и выключения коммутатора, число импульсов дребезга, и время удержания коммутатора во включенном состоянии. В рамках проведенного эксперимента значение TMW_CNT было принято значению эквивалентному 1 секунде. Был автоматизирован процесс замыкания коммутаторов с помощью Home Assistant Service. А также собрана статистика по распределению значений диапазоне от 0 до 255.

Данная схема является универсальным методом получения случайных последовательностей, так она не требует больших аппаратных затрат на реализацию и обладает хорошими статистическими свойствами вырабатываемых последовательностей. А также может быть использована как дополнительный источник энтропии в системах взаимодействующих с пользователем.

СПИСОК ЛИТЕРАТУРЫ

1. Можейко Д. О. Исследование дребезга контактов с целью генерирования случайных чисел / Д. О. Можейко, А. А. Иванюк // материалы 59-ой научной конференции аспирантов, магистрантов и студентов БГУИР, 2023 года, Минск, Республика Беларусь / – Минск: БГУИР, 2023. – С. 41-43
2. Maxfield M. Switch Bounce and Debounce (Part 1): Switch Types [Electronic resource] / M. Maxfield. — Mode of access: <https://www.eeweb.com/switch-bounce-and-debounce-part-1-switch-types>. – Date of access: 03.05.2023.
3. Переключатели электромеханические для электрического и электронного оборудования. Общие технические условия: Межгосударственный стандарт ГОСТ ИЕС 61020-1-2016. Ч. 1. – Москва: Стандартинформ, 2017. – 50 с.
4. Zybo Z7: Zynq-7000 ARM/FPGA SoC Development Board [Electronic resource]. – Mode of access: <https://digilent.com/reference/programmable-logic/zybo-z7/start>. – Date of access: 15.05.2023.

УСТРОЙСТВО ДЕТЕКТИРОВАНИЯ И ДЕКОДИРОВАНИЯ ДТМФ-СИГНАЛОВ

Осипов А. С., Вашкевич М. И.

Кафедра электронных вычислительных средств,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: cyklop3345@gmail.com

В работе представляется устройство детектирования и декодирования ДТМФ-сигналов на базе платы Zynq 7000. Данное устройство принимает на вход акустический сигнал и, в случае регистрации тонального набора, создает управляющее четырехразрядное слово.

ВВЕДЕНИЕ

DTMF (dual-tone multi-frequency) сигналы можно использовать для кодирования и передачи ограниченного набора из 16-ти символов. Часто передаваемый символ несет информацию об управляющем воздействии. Передача сигнала осуществляется по одному каналу связи, что обеспечивает связь между двумя устройствами с минимальными затратами средств.

Изначально тональный набор использовался в средствах телефонии для соединения аналогового оборудования, такого как телефонные аппараты и автоматические телефонные станции. Тональные сигналы применяются для систем голосового автоответа при ручном вводе команд абонентом.

На данный момент ДТМФ-сигналы обрели популярность и в других сферах помимо телефонии. Так в статье [8] предлагается использовать тональный набор для управления бытовой техникой в системе умного дома. А в работе [9] представлено удаленное управление роботом с использованием ДТМФ и GPS.

В данной статье предлагается использовать для связи аудиоканал, что позволит передать команду от управляющего устройства к одному или нескольким управляемым посредством звуковых сигналов. Помехи в аудиоканале, такие как человеческая речь или посторонние звуки, не будут накладываю ограничения на использование ДТМФ. Данная система может быть особенно востребована в местах с загруженными каналами Wi-Fi или Bluetooth в условиях невозможности проложить сигнальный кабель.

I. ФОРМИРОВАНИЕ ДТМФ-СИГНАЛА

ДТМФ-сигнала представляет собой сумму двух синусоид определенных частот. Частоты подобраны особым образом, чтобы обеспечить надежную передачу сигнала через телефонные линии и минимизировать возможные искажения и ошибки в распознавании сигнала.

Для формирования ДТМФ-сигналов используется восемь различных частот, которые разделены на группу нижних и группу верхних частот. Кодирование символов происходит путем сложения

двух синусоид (по одной из каждой группы) как показана в таблице 1.

Таблица 1 – Кодирование символов ДТМФ-сигналов

Нижняя группа, Гц	Верхняя группа, Гц			
	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

Рекомендуемая минимальная продолжительность ДТМФ-сигнала составляет 40 мс, а длительность пауз между сигналами не должна превышать 24 мс.

Подробная информация о выборе частот и параметров сигналов указана в общей рекомендации по телефонной коммутации и сигнализации [1-2].

II. АЛГОРИТМ ДЕКОДИРОВАНИЯ

Расшифровка тонального сигнала основана на алгоритме Герцеля [3], который представляет собой рекурсивный фильтр второго порядка. Фильтр может быть описан следующими выражениями:

$$y(-2) = y(-1) = 0;$$

$$y(n) = x(n) + \alpha \cdot y(n-1) - y(n-2);$$

$$A^2(n) = y^2(n) - \alpha \cdot y(n) \cdot y(n-1) + y^2(n-1).$$

$$\alpha = 2 \cdot \cos\left(\frac{2\pi f}{f_s}\right)$$

Значение A^2 – выходное значение фильтра, обозначающее наличие нужного тона, а также квадрат его амплитуды.

Изначальный алгоритм Герцеля имеет следующие недостатки:

- на выходе фильтра сигнал колеблется от 0 до максимальной амплитуды, что не позволяет правильно считать амплитуду в произвольный такт работы фильтра;
- после исчезновения тонального набора на входе фильтра, сигнал на выходе остается вплоть до отключения фильтра;
- сигнал на выходе ненормирован.

Для решения некоторых недостатком можно разделить работу фильтра на периоды детектирования и принудительной очистки буфера. Но

данная мера потребуют дополнительной синхронизации работы передатчика и приемника.

В данном устройстве для исправления представленных недостатков была применена следующая модернизация алгоритма:

- в рекурсивной части фильтра вводится принудительное затухание сигналов;
- на выходе фильтра сигнал нормируется;
- после нормирования сигнал обрабатывается по методу угасающего максимума и по пороговой схеме для получения значения 0 (тон отсутствует) либо 1 (тон присутствует) на выходе.

Примененные методы позволяют непрерывно декодировать сигналы любой длины в любой момент времени.

Более полное описание алгоритма Герцеля можно найти в работах [4–7].

III. РЕАЛИЗАЦИЯ НА ПЛИС

Для реализации устройства была выбрана отладочная плата Zybo на базе ПЛИС Zynq-7000. Данная плата обладает разъемом под микрофон и четырьмя светодиодами, которые могут быть использованы для проверки правильности работы устройства.

На рисунке 1 представлена структурная схема устройства детектирования DTMF-сигналов на базе Zynq 7000. Ввод осуществляется посредством подачи тональных сигналов на микрофон, который через разъем TRS (mini-jack 3.5mm) соединяется с кодеком SSM2603. Настройка кодека осуществляется с использованием процессорного ядра ARM посредством интерфейса I²S. Аудиоданные от кодека передаются в процессорное ядро ARM по интерфейсу I²S и затем перенаправляются через AXI интерфейс в детектор DTMF сигналов, из которого результат детектирования выводится на LED-индикаторы.

Описание детектора DTMF сигналов выполнено на языке VHDL. Устройство состоит из восьми фильтров, подключенных параллельно, и схемы принятия решений. Единовременно тон могут

регистрировать только по одному фильтру из верхней и нижней группы. Таким образом, на выходе образуется два кода «1 из 4». Схема принятия решений преобразует получившиеся коды в код 8421. Помимо этого в схеме присутствуют защитные меры по фильтрации «ошибочных» комбинаций и фильтрации коротких сигналов, вызванных дребезгом сигналов.

IV. ЗАКЛЮЧЕНИЕ

В работе предложен вариант реализации устройства детектирования/декодирования DTMF-сигналов на базе платформы Zybo. Вычислительное ядро детектора DTMF-сигнала предложено реализовать в виде IP-блока с использованием модифицированного алгоритма Герцеля.

1. ITU-T Recommendation Q.23: Technical features of push-button telephone / 1993
2. ITU-T Recommendation Q.24: Multi-frequency push-button signal reception / 1993
3. Kazus [Электронный ресурс]. – Режим доступа : <https://kazus.ru/articles/149.html>.
4. Design of DTMF signal detection method based on improved Goertzel algorithm / X. Peiyao, L. Xiaofang, L. Weican, L. Mingyu, 2023 – Mode of access: <https://ieeexplore.ieee.org/document/10075142>
5. Improving the Goertzel-Blahut Algorithm – S. V. Fedorenko, 2016 – Mode of access: <https://ieeexplore.ieee.org/document/7457627>
6. Design of DTMF signal experiment system based on MATLAB – J. Long, 2020 – Mode of access: <https://ieeexplore.ieee.org/document/9315651>
7. The FPGA Implementation of Modified Goertzel Algorithm for DTMF Signal Detection – Z. Xinyi, 2010 – Mode of access: <https://ieeexplore.ieee.org/document/5630500>
8. Design of Remote Control System for Household Appliances based on Single Chip Microcomputer – H. Haibo, 2020 IOP Conf. Ser.: Mater. Sci. Eng. 750 012109 – Mode of access: <https://iopscience.iop.org/article/10.1088/1757-899X/750/1/012109>.
9. Development of Real Time Night Vision Camera Monitoring Robot Integrating DTMF and GPS System / M. S. Sulong, M. Z. Hasan, N. Shariffudin, M. A. Busari and M. N. Mansor – et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 917 012033 – Mode of access: <https://iopscience.iop.org/article/10.1088/1757-899X/917/1/012033>

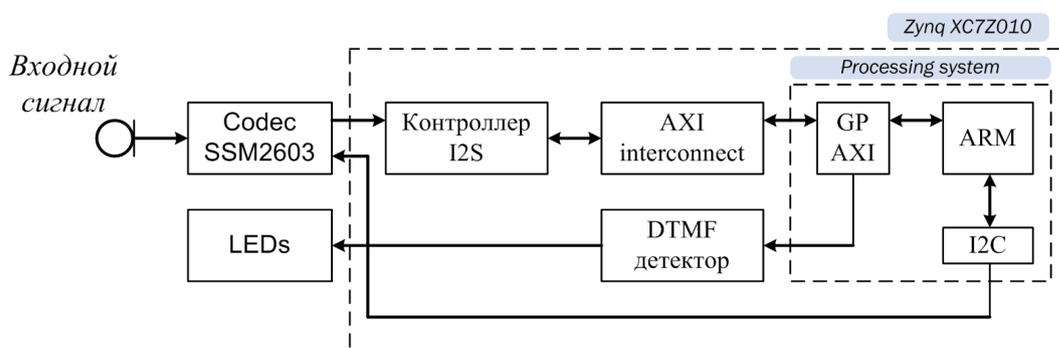


Рис. 1 – Структурная схема устройства детектирования DTMF-сигналов на базе платформы Zybo

СИСТЕМА НА КРИСТАЛЛЕ ДЛЯ АППАРАТНО-ПРОГРАММНОЙ ОБРАБОТКИ ЧАСТОТНО-МОДУЛИРУЕМЫХ СИГНАЛОВ

Силантьев А. М., Примаков Е. В., Переверзев А. Л., Осадчий И. С.
Институт МПСУ, Национальный исследовательский университет
«Московский институт электронной техники»
Зеленоград, Российская Федерация
E-mail: silantiev@org.miet.ru

На основе анализа структуры радиолокационных датчиков малой дальности с использованием непрерывного частотно-модулированного излучаемого сигнала предложена его реализация с помощью специализированной системы на кристалле на архитектуре RISC-V. Рассмотрены характеристики, принципы построения и основные алгоритмы обработки сигналов.

ВВЕДЕНИЕ

Радиолокационные системы (РЛС) широко применяются в различных военных и гражданских отраслях, например, в авиации, морском и наземном транспорте, метеорологии и даже в археологии. Особое место занимают РЛС малой дальности, ориентировочно, до десятков метров, которые применяются, в частности, в беспилотных летательных аппаратах (БПЛА) и автомобилестроении.

В БПЛА набор радиолокационных датчиков может быть использован для точной оценки высоты аппарата над землёй, его автоматизированной посадки, а также для обнаружения и избегания препятствий при полёте. В автомобилях датчики РЛС малой дальности используются как составная часть систем помощи водителю (ADAS - Advanced Driver-Assistance Systems) [1]. Возможно применение РЛС малой дальности действия для определения характеристик дорожного полотна [2] и работы адаптивной системы подвески. В МИЭТ также создана и экспериментально опробована бесконтактная радиолокационная система измерения толщины льда перед автомобилем [3] для безопасного передвижения по зимникам в руслах замерзших рек, что особенно важно для практически лишенных дорог зон севера РФ и Сибири.

I. ОСНОВНАЯ ЧАСТЬ

В НИУ МИЭТ накоплен существенный опыт по проектированию подобных систем для различных применений в С-, X- и К- диапазонах частот. Вопрос уменьшения стоимости радиолокационных датчиков и их доступности для потребителя становится ключевым. Одним из путей удешевления аппаратуры РЛС малой дальности является сокращение производственных издержек за счет уменьшения номенклатуры и количества применяемых элементов, создания специализированных наборов интегральных микросхем, исключения процедур настройки и проверки отдельных узлов устройства.

Наиболее распространенным и простым решением технической реализации РЛС малой дальности действия является использование непрерывного излучаемого сигнала с частотной модуляцией (Frequency Modulated Continuous Wave – FMCW) [4], [5]. Обобщенная блок-схема FMCW радиолокационного датчика приведена на рис.1. Датчик включает в себя два основных блока – приемно-передатчик и устройство цифровой обработки сигналов (ЦОС), а также приемную и передающую антенны.

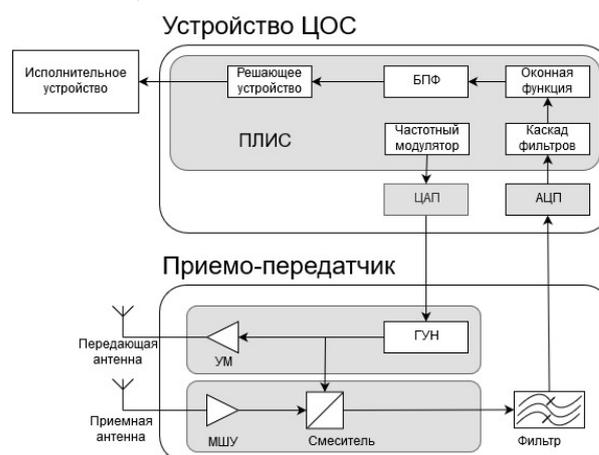


Рис. 1 – Блок-схема РЛС ближней зоны с непрерывным ЧМ излучением

Передающий канал образован генератором управляемым напряжением (ГУН) и усилителем мощности (УМ). Приемный канал содержит малошумящий усилитель (МШУ) и смеситель, использующий в качестве опорного напряжения тот же сигнал, который излучается передатчиком. Таким образом, в приемно-передатчике происходит преобразование на нулевую промежуточную частоту. Выходной фильтр приемного канала выделяет требуемый диапазон частот биений.

Устройство ЦОС, включает в себя аналого-цифровой (АЦП) и цифроаналоговый (ЦАП) преобразователи, а также цифровой процессор. Устройство ЦОС отвечает за формирование сигнала управления частотой передатчика, который через ЦАП поступает на вход управления частоты

той ГУН. Выходной сигнал приемного канала оцифровывается с помощью АЦП и поступает в цифровой процессор на базе программируемой логической интегральной схемы (ПЛИС). Подобная структура позволяет строить радиолокационные датчики доплеровского, дальномерного и комбинированного типа, причем конкретное функционирование определяется только программой ЦОС.

Для решения описанных выше проблем создается специализированная система на кристалле (СнК) (рис.2) основанная на процессорном ядре с открытой архитектурой RISC-V, ЦОС-ускорителе и аналогово-цифровой обвязке в виде ЦАП, АЦП, памяти.

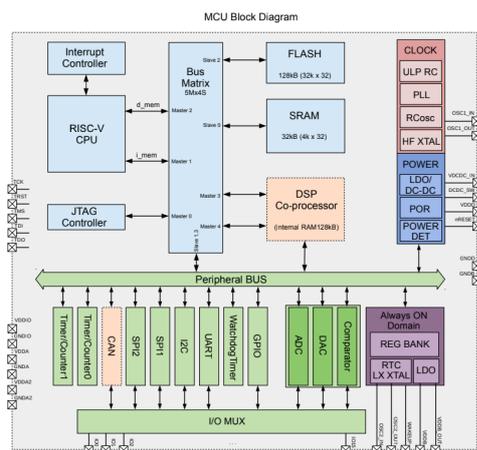


Рис. 2 – Блок-схема микросхемы ЦОС

Алгоритм работы ЦОС разбивается на первичную и вторичную обработку данных, которые отличаются, главным образом, скоростью обработки и объемом потока данных. Первичная обработка данных включает в себя операции над высокоскоростным потоком входных данных: первичную фильтрацию, взвешивание и анализ частотного спектра сигнала на основе алгоритма быстрого преобразования Фурье, адаптивное пороговое обнаружение сигнала. Для выполнения первичной обработки в масштабе реального времени требуется высокая производительность вычислительного устройства. Для повышения производительности при сохранении низкого энергопотребления в состав системы включены специализированные аппаратные блоки-ускорители.

В задачи вторичной обработки входит изменение параметров объекта наблюдения (дальности, скорости, траектории), принятие решения на срабатывание исполнительных устройств. Методы вторичной обработки могут различаться как для разных применений, так и для различных режимов работы одного устройства. Исходя из

этого, наиболее эффективно выполнять вторичную обработку сигнала программно, используя процессорное ядро.

Таким образом, СнК включает в себя, как специализированное высокопроизводительное ядро, предназначенное для первичной обработки сигнала, так и 32-разрядное процессорное ядро для вторичной обработки, управления и обмена данными по внешним интерфейсам. Ключевыми особенностями СнК являются высокая производительность, низкое энергопотребление и занимаемая площадь на кристалле, возможность гибкого конфигурирования для использования в составе разных по характеристикам и назначению устройств. Прототип СнК произведен по технологии 180 нм и обладает следующими характеристиками: тактовая частота пяти-стадийного RISC-V процессора до 120 МГц, размер ОЗУ 128 Кбайт, BootROM с загрузкой из внешней QSPI-памяти, интерфейсы CAN, SPI, I2C, GPIO, UART, наличие параллельных интерфейсов для внешних АЦП и ЦАП, интегрированный АЦП 12-бит с скоростью до 15 MSPS, интегрированный ЦАП 12-бит с скоростью до 5 MSPS, тактовая частота интегрированного ЦОС-ускорителя до 100 МГц. В ЦОС-ускорителе реализованы конфигурируемые потоковые вычислители для фильтра постоянной составляющей, КИХ-фильтра 128 порядка, оконной функции, СС-фильтра, конвейерного БПФ на 4096 точек, вычисление модуля комплексного числа, DMA-блок с детектированием пиков в спектре и блок управления модуляцией и приемом информационного сигнала.¹

СПИСОК ЛИТЕРАТУРЫ

1. Advanced driver-assistance systems [Электронный ресурс] Режим доступа: https://en.wikipedia.org/wiki/Advanced_driver-assistance_systems Дата обращения: 14.10.2023.
2. A. Bystrov, E. Hoar, T.-Y. Tran, N. Clarke, M. Gashinova, M. Cherniakov. Sensors for Automotive Remote Road Surface Classification, – Proc. 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Sept. 12-14, 2018, Madrid, Spain.
3. Ju. Kosova, D. Kaleev, V. Sizov, P. Metelkov, Device for Determining the Thickness and Properties of Ice Cover on the Water Surface Using a Frequency-Modulated Signal. – Proc. 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus-2019), Moscow, MIET, 2019, pp. 1639 – 1643.
4. И.В. Комаров, С.М. Смольский, "Основы теории радиолокационных систем с непрерывным излучением частотно-модулированных колебаний". – М.: Горячая линия–Телеком, 2010. – 366 с.
5. G. M. Brooker, Understanding Millimetre Wave FMCW Radars, Proc. 1st International Conference on Sensing Technology, Nov. 21-23, 2005 Palmerston North, New Zealand, pp. 152-157.

¹Производство интегральной микросхемы было выполнено за счет средств Минобрнауки России в рамках федерального проекта "Подготовка кадров и научного фундамента для электронной промышленности" по гос.заданию на выполнение научно- исследовательской работы "Разработка методики прототипирования электронной компонентной базы на отечественных микроэлектронных производствах на основе сервиса MPW (FSMR-2023-0008)".

РАСПОЗНАВАНИЕ ЛОГИЧЕСКИХ СТРУКТУР НА ОСНОВЕ ПРОХОДНЫХ ТРАНЗИСТОРОВ В КМОП СХЕМЕ

Черемисинов Д. И., Черемисинова Л. Д.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: {cher, cld}@newman.bas-net.by

Рассматривается задача распознавания в КМОП схеме подсем из проходных транзисторов, являющихся передаточными элементами и мультиплексорами (демультиплексорами).

ВВЕДЕНИЕ

Современные СБИС содержат сотни миллионов транзисторов и связей между ними. Важным инструментом автоматизированного проектирования схем такой сложности являются средства распознавания высокоуровневых структур в схемах на транзисторном уровне. Построение иерархического структурного описания по плоскому (одноуровневому) структурному описанию является существенной операцией в процессе верификации, проектирования и перепроектирования интегральных схем.

В статье [1] описан метод и программа декомпиляции КМОП схемы из транзисторов, основным компонентом которой является нахождение подсем из транзисторов, реализующих КМОП вентили. В настоящей работе предлагаются методы (и их программная реализация) для выявления в КМОП схеме, заданной плоским нетлстом в формате SPICE (Simulation Program with Integrated Circuit Emphasis), подсем проходной транзисторной логики (PTL – pass transistor logic), представляющих собой передаточные элементы и мультиплексоры (демультиплексоры).

I. ЭЛЕМЕНТЫ КМОП ЛОГИКИ НА ОСНОВЕ ПРОХОДНЫХ ТРАНЗИСТОРОВ

Логические элементы на основе стандартной КМОП логики состоят из равномогных групп р-МОП и n-МОП транзисторов, их недостатком является значительное количество МОП транзисторов, в частности р-МОП. В КМОП схемах наряду с КМОП вентилями часто используются логические элементы на основе цепей из проходных транзисторов.

На выводы проходных р-МОП и n-МОП транзисторов подаются входные и внутренние сигналы схемы (вместо сигналов питания и земли). n-МОП транзистор является почти идеальным переключателем при передаче сигнала логического 0, но сильно снижает уровень логического сигнала 1. р-МОП транзистор, наоборот, почти идеально пропускает сигнал 1, но не совершенен при передаче сигнала 0. Для устранения этих недостатков в качестве КМОП ключа применяют передаточный вентиль – пару параллельно соединенных n-МОП и р-МОП транзисторов (рис. 1), для которых используется парафазное управление: на

вход р-МОП транзистора подается управляющий сигнал C , а на вход n-МОП инвертированный – \bar{C} . Передача сигнала в таком вентиле может идти в обоих направлениях под управлением напряжения C и \bar{C} на затворах МОП транзисторов.

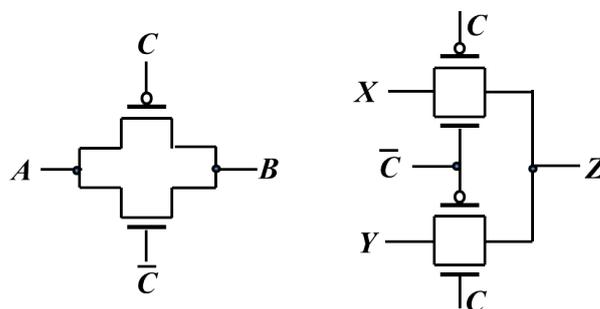


Рис. 1 – Передаточный вентиль и мультиплексор на его основе

Передаточные вентили широко используются при проектировании регулярных структур, таких как мультиплексоры, демультиплексоры, триггеры, ПЗУ, триггеры.

Мультиплексор представляет собой устройство, которое позволяют подключать несколько входов к одному выходу. КМОП мультиплексор на основе проходной логики пропускает ток в обоих направлениях, позволяя коммутировать не только цифровые, но и аналоговые сигналы. На рис. 1 приведена схема мультиплексора (селектора) 2×1 , имеющего два входа X и Y и один выход Z . Сигналы X и Y подаются на передаточные элементы $G1$ и $G2$, селектирующий сигнал C управляет передачей на выход одного из сигналов X или Y . $C = 0$ открывает элементы $G1$, пропуская на выход сигнал X . Если $C = 1$, то на выход проходит сигнал Y . Выходная функция мультиплексора $Z = X\bar{C} \vee YC$.

Схема, представленная на рис. 1, может выполнять и обратную функцию – функцию демультиплексора 1×2 , передавая сигнал с выхода Z на входы X и Y . В этом случае сигнал с входа Z поступает на выход X , когда $C = 0$, и на выход Y , когда $C = 1$. Функции X и Y демультиплексора могут быть записаны как $X = Z\bar{C}$ и $Y = ZC$.

II. ЗАДАНИЕ ТРАНЗИСТОРНЫХ СХЕМ НА

Исходная плоская и результирующая иерархическая транзисторные схемы представляются в формате проектирования SPICE. Главной частью SPICE-описания схемы является список транзисторов, в котором для каждого вывода транзистора (сток, затвор, исток, подложка) указано имя цепи, соединяющей его с остальными частями схемы. Описание связей униполярного транзистора в формате SPICE имеет вид:

M<name> <nd> <ng> <ns> <nb> <model-name> ,

где name – название транзистора; nd, ng, ns и nb – идентификаторы цепей, связанных с выводами стока (drain), затвора (gate), истока (source) и подложки (substrate) соответственно; model-name – тип транзистора: n-МОП или p-МОП (nmos или pmos). Например, SPICE-описание CN2 передаточного элемента (рис. 1), состоящего из двух транзисторов, имеет вид (здесь цепь nC соответствует сигналу \bar{C}):

```
.SUBCKT CN2 A nC B C
M1 A C B vcc pmos
M2 A nC B gnd nmos
.ENDS
```

III. РАСПОЗНАВАНИЕ ПЕРЕДАТОЧНЫХ ВЕНТИЛЕЙ

Передаточные элементы распознаются путем поиска пар n-МОП и p-МОП транзисторов, соединённых параллельно выводами стока и истока. На следующем листинге выделены две пары транзисторов, составляющие передаточные вентили:

```
.SUBCKT rdrv1 r1 s1 c d vcc gnd q
M0 130 8 9 gnd MN
M1 gnd r1 130 gnd MN
M2 131 s1 gnd gnd MN
M3 13 g 131 gnd MN
M4 8 c 13 gnd MN
M5 d 10 8 gnd MN
M6 gnd c 10 gnd MN

M18 13 g vcc vcc MP
M19 8 10 13 vcc MP
M20 d c 8 vcc MP
M21 vcc c 10 vcc MP
M22 12 10 g vcc MP
```

Временная сложность поиска связанных транзисторов путем попарного сравнения строк SPICE описания пропорциональна числу сочетаний из n (число транзисторов в схеме) по 2. Для реальных схем, содержащих миллионы транзисторов, это не приемлемо.

В работе предлагается метод поиска параллельно связанных транзисторов линейной сложности с помощью хеш-таблицы. Таблица строится в процессе последовательного анализа строк исходного SPICE-описания, когда находится графовая модель схемы. Ключами хеш-таблицы являются хеши, вычисленные по текстовым строкам, образованным из имен цепей истока и стока

каждого из транзисторов анализируемой схемы. Ситуация, когда для очередного рассматриваемого ключа в хеш-таблице уже есть вход, означает, что транзисторы с этим хешем составляют передаточный вентиль.

IV. РАСПОЗНАВАНИЕ МУЛЬТИПЛЕКСОРОВ

Распознавание мультиплексоров 2×1 осуществляется на этапе анализа схемы, когда передаточные вентили уже распознаны. Тогда для них строится хеш-таблица. Ключами этой таблицы служат хеши от имен цепей, подключенных к выводам A и B каждого из передаточных элементов (модели CN2). Для каждого элемента вычисляется два хеша: для A, B и B, A. Таким образом, каждый вентиль попадает в две записи хеш-таблицы. С точки зрения топологии не имеет значения, какую функцию (мультиплексирования или демультиплексирования) выполняет распознанная схема. По умолчанию считается, что это схема работает как мультиплексор.

Ситуация, когда для очередного рассматриваемого ключа в хеш-таблице уже есть вход, означает, что существует два передаточных элемента G1 и G2, которые потенциально могут быть составлять мультиплексор. Для таких пар элементов производится проверка фактических параметров цепей C и nC, связанных с затворами транзисторов: необходимо, чтобы был взаимно обратный порядок их следования в этих элементах.

На следующем листинге:

```
.SUBCKT fdrv1s1 r1 s1 c d q
X0I1 r1 11 14 G0_0
X0I2 8 r1 9 G0_1
X0I3 s1 12 11 G0_2
X0I4 s1 9 13 G0_3
X1I1 11 qn G1_0
X1I2 c 10 G1_0
X1I3 12 q G1_1
X2I1 13 10 8 c CN2
X2I2 8 c d 10 CN2
X2I3 9 10 12 c CN2
X2I4 12 c 14 10 CN2
.ENDS
```

элементы X2I1 и X2I2 попадают в одну запись хеш-таблицы, так как $A_2 = B_1 = 8$ и, кроме того, $nC_1 = 10$, $C_1 = c$, а $nC_2 = c$, $C_2 = 10$. Таким образом, элементы X2I1 и X2I2 составляют мультиплексор. Аналогично элементы X2I3 и X2I4 также образуют мультиплексор.

V. ЗАКЛЮЧЕНИЕ

Предложенные методы распознавания передаточных вентилях и мультиплексоров в КМОП схеме реализованы на языке C++ и включены в программу декомпиляции плоской транзисторной схемы [1].

СПИСОК ЛИТЕРАТУРЫ

1. Черемисинов, Д.И. Извлечение сети логических элементов из КМОП схемы транзисторного уровня / Д.И. Черемисинов, Л.Д. Черемисинова // Микроэлектроника. – № 3 (48). – 2019. – С. 224–234.

ТРАНСФОРМАЦИЯ ИЕРАРХИЧЕСКОГО ОПИСАНИЯ КМОП СХЕМЫ В ПЛОСКОЕ ПРЕДСТАВЛЕНИЕ НА ЯЗЫКЕ SPICE

Черемисинова Л. Д., Черемисинов Д. И.

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: {cld, cher}@newman.bas-net.by

Рассматривается задача преобразования иерархических структурных описаний транзисторных КМОП схем, представленных на языке SPICE в одноуровневые описания. Предлагается рекурсивный метод компиляции иерархических описаний и его программная реализация на языке C++.

ВВЕДЕНИЕ

Современные цифровые КМОП схемы содержат до миллиарда примитивных элементов на транзисторном уровне. Ответственным этапом проектирования таких схем является тестирование результатов проектирования, которое позволяет установить, соответствует ли описание проекта на уровне транзисторов спецификации проектируемого устройства. Важными инструментами, позволяющими снизить трудоемкость тестирования электрических схем такой сложности, являются средства декомпиляции и компиляции описаний транзисторных схем. Декомпиляция позволяет, исходя из плоского (одноуровневого) описания транзисторной схемы, восстановить иерархическое структурное описание на уровне логических элементов [1]. Компиляция, наоборот, применяется к иерархическому структурному описанию схемы и позволяет преобразовать его в плоское описание. Следует заметить, что компиляция и декомпиляция изменяют не саму схему, а только ее представление.

В настоящей работе рассматривается задача компиляции иерархических описаний транзисторных схем для наиболее распространенного стиля логики – логических комплементарных МОП-структур. Исходное и полученное в результате компиляции описания транзисторной схемы задаются в формате SPICE (Simulation Program with Integrated Circuit Emphasis).

I. ЗАДАНИЕ ТРАНЗИСТОРНЫХ СХЕМ НА ЯЗЫКЕ SPICE

Главной частью описания (для целей компиляции) схемы в формате SPICE является раздел Data Statements, который описывает компоненты транзисторной схемы и взаимосвязи между ними. Компонентами могут быть примитивные элементы – транзисторы – и элементы блочного типа, которые представляются подсхемами.

Общая форма описания связей униполярного транзистора в формате SPICE имеет вид

M<name> <nd> <ng> <ns> <nb> <model-name> ,

где «name» – имя транзистора; «nd», «ng», «ns» и «nb» – идентификаторы цепей, связанных с

выводами транзистора: стока (drain), затвора (gate), истока (source) и подложки (substrate) соответственно; «model-name» – имя модели – типа транзистора. Пример SPICE-описания КМОП инвертора, состоящего из двух MOSFET транзисторов приведен на рис. 1.

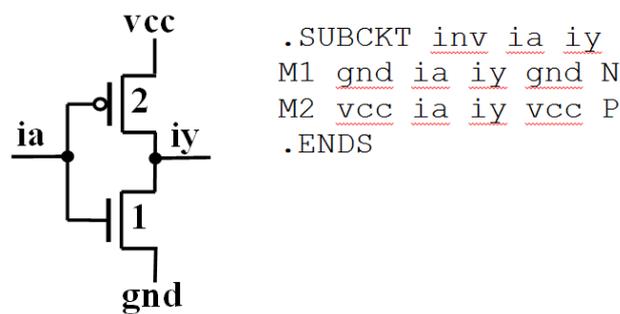


Рис. 1 – Схема инвертора inv и ее SPICE описание

Общая форма описания связей непримитивного элемента с n выводами имеет вид:

X<name> <P1> <P2> ... <Pn> <model-name> ,

где «name» – имя элемента; «P1», «P2»... «Pn» – идентификаторы цепей, связанных с выводами элемента; «model-name» – имя модели – типа элемента.

На следующем листинге ниже приведено описание простой КМОП схемы в виде трехуровневой иерархической структуры (верхняя часть рис. 2):

```
* SPICE deck for cell data bufer
.GLOBAL vcc gnd
.SUBCKT inv ia iy
M1 gnd ia iy gnd N w=2.4e-06 l=1e-06
M2 vcc ia iy vcc P w=3.2e-06 l=1e-06
.ENDS
.SUBCKT buf bA bY
Xin1 bA b1 inv
Xin2 b1 bY inv
.ENDS
.SUBCKT data bufer A Y
Xb1 A Y buf
.ENDS
```

Схема верхнего уровня – *data-bufer* – содержит один экземпляр схемы *buf*. Схема *buf*, в свою очередь, состоит из двух последовательно соединенных схем инверторов типа *inv* (рис. 1).

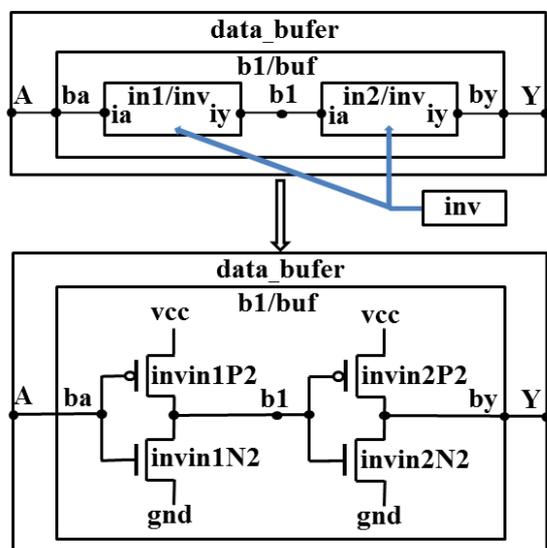


Рис. 2 – Результат преобразования схемы после компиляции элементов типа *inv*

II. ОБЩАЯ СХЕМА КОМПИЛЯЦИИ ОПИСАНИЯ ИЕРАРХИЧЕСКОЙ СХЕМЫ И В ПЛОСКОЕ ПРЕДСТАВЛЕНИЕ

Процесс компиляции состоит из трех шагов. На первом шаге выполняется анализ исходного описания схемы в формате SPICE: собирается вся информация, связанная с каждым синтаксическим элементом схемы, и строятся внутренние графовые представления каждого нетривиального элемента. На втором этапе производится собственно компиляция иерархического описания, начиная с элемента верхнего уровня иерархии. Граф схемы просматривается для обнаружения экземпляров нетривиальных элементов, каждый из которых заменяется его схемой. При этой замене исключаются порты подсхем компилируемых элементов. На последнем шаге выполняется сериализация внутреннего плоского представления графа скомпилированной схемы в формат SPICE.

Процедура компиляции схемы является рекурсивной. Если в схеме элемента, компилируемого в настоящий момент, встречается непримитивный элемент (его имя начинается с «X»), то запускается процедура его компиляции, и так до достижения листовых элементов в иерархической цепочке. При этом, результаты компиляции заменяют исходные иерархические описания

моделей элементов, и если в схеме встречаются элементы, имеющие тип, соответствующий уже ранее скомпилированной схеме, то повторная компиляция не требуется.

Так как выводы экземпляров элементов одного и того же типа в материнских схемах могут быть подключены к разным цепям, то цепи, соответствующие портам описания модели этих элементов, переименовываются, получая имена цепей, являющихся параметрами конкретного экземпляра элемента в рассматриваемой схеме. Процесс такой замены аналогичен принятой в программировании подстановке фактических параметров вместо формальных при вызове подпрограммы. Кроме того, изменяются также и имена элементов в самой схеме модели, подставляемой вместо компилируемого экземпляра элемента: они снабжаются префиксом, представляющим собой путь в иерархии до компилируемого элемента. Необходимость такого переименования вызвана тем, что в компилируемой схеме может не один элемент одного и того же типа, что приводит к повторению схемы модели в скомпилированной схеме.

На рис. 2 показан результат преобразования схемы, заданной на выше приведенном листинге, после выполнения компиляции модели *buf* путем замены экземпляров элементов *in1* и *in2* схемой модели *inv*. Плоское описание схемы *data-bufer*, полученное в результате ее компиляции, приведено на следующем листинге.

```
* SPICE deck for cell data_bufer
.GLOBAL gnd vcc
.SUBCKT data_bufer A Y
Mbufb1/invin1/N1 gnd A bufb1/b1 gnd N
Mbufb1/invin1/P2 vcc A bufb1/b1 vcc P
Mbufb1/invin2/N1 gnd bufb1/b 1 Y gnd N
Mbufb1/invin2/P2 vcc bufb1/b1 Y vcc P
.ENDS
```

III. ЗАКЛЮЧЕНИЕ

Процедура компиляции реализована на языке C++ и включена в программу анализа транзисторных схем на эквивалентность, которая была протестирована на ряде практических примеров. Программа имеет достаточное быстродействие, чтобы обрабатывать схемы с более чем 500 тысячами транзисторов за несколько минут на персональной ЭВМ.

СПИСОК ЛИТЕРАТУРЫ

1. Черемисинов, Д.И. Извлечение сети логических элементов из КМОП схемы транзисторного уровня / Д.И. Черемисинов, Л.Д. Черемисинова // Микроэлектроника. – № 3 (48). – 2019. – С. 224–234.

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ДЛЯ БЫСТРОГО ПРОТОТИПИРОВАНИЯ И ИССЛЕДОВАНИЯ КРИПТОГРАФИЧЕСКИХ ПРИМИТИВОВ НА БАЗЕ ПЛИС

Шамына А. Ю., Кайкы М. Н., Иванюк А. А.

Кафедра программного обеспечения информационных технологий, кафедра информатики,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: shamyna@bsuir.by, kaikymykhailo@gmail.com, ivaniuk@bsuir.by

В настоящей работе описываются построение, инструментарий и конфигурация аппаратно-программного комплекса, предназначенного для проведения исследований криптографических примитивов, реализованных на FPGA в составе плат быстрого прототипирования. Рассмотрены нюансы построения сетевой, аппаратной и программной части исследовательского комплекса.

ВВЕДЕНИЕ

В настоящее время использование средств физической криптографии приобретает все более широкий характер [1]. Одновременно с этим растет интерес к исследованиям в данной области. Зачастую разработчики и исследователи в качестве платформы для реализации различных схем физической криптографии выбирают FPGA. Однако, наиболее трудоемким этапом подготовки эксперимента является конфигурация инфраструктуры для сбора и анализа данных. Кроме этого, существует проблема проведения экспериментов при стационарных внешних условиях (например, температура, напряжение питания и т.п.). Немаловажной является проблема представления и анализа данных экспериментов в максимально унифицированном виде. Одной из главных целей разработки аппаратно-программного комплекса является стремление решить эти проблемы и предоставить возможность исследователям максимально сократить время развертывания эксперимента.

I. ОПИСАНИЕ УСТАНОВКИ

Центральным узлом комплекса является хост – ПЭВМ на базе CPU Intel Xeon Bronze 3104 с 32 ГБ ОЗУ под управлением ОС Ubuntu 22.04. В состав ПЭВМ входят также несколько сетевых карт, где одна из них подключена к сети лаборатории, а вторая - к сетевому коммутатору с подключенными к нему платами быстрого прототипирования Digilent Zybo Z7-10 [2] (см. рис. 1). Для обеспечения стационарности температуры используется температурная камера TestEquity 155 Benchtop [3] с диапазоном температур от -20 °C до +130 °C и возможностью удаленного управления с использованием протокола SCPI (Standard Commands for Programmable Instruments). В качестве управляемого лабораторного блока питания выбран ELEMENT 305DB. В состав выбранной модели платы быстрого прототипирования входит ЧИК Xilinx Zynq 7000 [4] с двоядерным процессором Cortex-A9 с частотой 667 МГц и 1 ГБ DDR3 RAM, что делает возможным работу ОС

Linux на данной платформе. В качестве дистрибутива Linux был выбран PetaLinux 2022.1 [5] с интегрированным клиентом синхронизации времени по протоколу NTP (Network Time Protocol) и драйвером, необходимым для загрузки конфигурации FPGA. Также был внедрен сервер TCF (Target Communications Framework) для сохранения возможности отладки программ из Vitis при работе PetaLinux. Образ ОС на плате загружается с карты памяти microSD. Для автоматического назначения IP-адресов сетевым интерфейсам плат на хосте был развернут DHCP-сервер с резервированием адресов из пула адресов по MAC-адресу сетевой карты. Для удобства управления платами на локальном DNS-сервере хоста были созданы DNS записи типа A, где имя для платы генерируется с использованием MAC-адреса. Сценарии работы и управления были написаны с использованием Bash. Управление платами осуществляется по протоколу SSH.

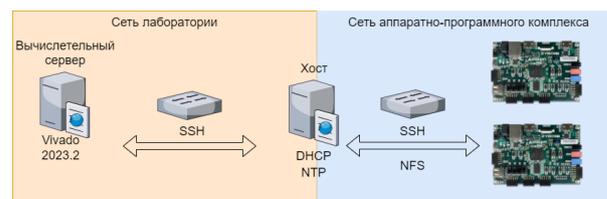


Рис. 1 – Схема аппаратно-программного комплекса

На хосте создана сетевая директория, в которой для каждой платы существует собственная поддиректория и которая включает в себя папки данных и скриптов. Эти папки являются сетевыми и монтируются в файловую систему отладочных плат в качестве NFS (Network File System) директорий.

II. ПОРЯДОК ВЫПОЛНЕНИЯ ЭКСПЕРИМЕНТА

Для проведения экспериментов необходимо подготовить IP-компонент исследуемого криптографического примитива, интерфейс которого должен быть выполнен по шаблону. HDL-описание аппаратуры содержит достаточное количество конфигурационных регистров и данных

для обеспечения работы исследуемых криптографических примитивов, которые спроецированы в адресное пространство PS (Processing System). Подключение компонентов, управляемых из адресного пространства PS осуществляется через интерфейс AXI4-LITE.

Затем подготовленный компонент подключается к проекту с HDL-описанием аппаратуры и уже вновь сгенерированный проект помещается по заданному пути.

На следующем этапе выполняется настройка скрипта Bash выполнения эксперимента, которая заключается в указании набора плат, на котором будет запускаться эксперимент, а также исходного кода программы на языке C, запускаемой на платах для опроса примитива. Затем выполняется запуск скрипта анализа данных экспериментов на Python. Есть также возможность указания количества экспериментов и реконфигураций блока программируемой логики (англ. Programmable Logic, PL), температурного режима и уровня напряжения питания плат. Алгоритм проведения эксперимента можно представить в виде последовательности действий (см. рис. 2).



Рис. 2 – Алгоритм выполнения эксперимента

После выполнения анализа, выполняется перемещение данных эксперимента на сетевой накопитель с указанием идентификатора эксперимента и временной отметки. Генерирование .bit файла конфигурации PL и синтез проекта осуществляется с использованием мощностей вычислительного сервера. Загрузка конфигурации FPGA производится из PS Zynq 7000 с использованием fpga_util без перезагрузки PetaLinux. Использование такого подхода для конфигурации PL требует генерирование из файла конфигурации .bit файла .bin, которое производится утилитой BitGen, доступной в пакете PetaLinux. Кроме полной реконфигурации PL доступно также ее частичная реконфигурация (англ. Partial Reconfiguration, PR) [6].

III. ЗАКЛЮЧЕНИЕ

В результате выполнения описанной работы был создан аппаратно-программный комплекс, который существенно снижает временные и трудовые затраты при выполнении исследований криптографических примитивов. Параметризация сценариев работы экспериментальной установки позволяет гибко ее настраивать под конкретный эксперимент. Использование высокоскоростной сетевой инфраструктуры 1 Гбит/с в сочетании с сетевыми хранилищами общим объемом 24 ТБ, построенных на RAID-массивах, позволяет работать с большим объемом экспериментальных данных. Потенциально данная схема экспериментов может быть расширена для исследования любых других схем физической криптографии.

СПИСОК ЛИТЕРАТУРЫ

1. Ярмолик, В. Н. Физически неклонируемые функции / В. Н. Ярмолик, Ю. Г. Вашинко // Информатика. – 2011. – №2. – С. 90-100.
2. Zybo Z7 Refence / Digilent inc [Electronic resource]. – Mode of access: <https://digilent.com/reference/programmable-logic/zybo-z7/start>. – Date of access: 10.10.2023.
3. TestEquity Chambers TE-155 Benchtop Temperature Chamber / TestEquity LLC [Electronic resource]. – Mode of access: <https://www.testequity.com/product/20200-1-TE-155>. – Date of access: 10.10.2023.
4. Zynq 7000 SoC / Advanced Micro Devices, Inc. [Electronic resource]. – Mode of access: <https://www.xilinx.com/products/silicon-devices/soc/zynq-7000.html>. – Date of access: 10.10.2023.
5. PetaLinux Tools Documentation: Reference Guide / Advanced Micro Devices, Inc. [Electronic resource]. – Mode of access: <https://docs.xilinx.com/r/2022.1-English/ug1144-petalinux-tools-reference-guide>. – Date of access: 10.10.2023.
6. Vivado Design Suite User Guide Dynamic Function eXchange / Advanced Micro Devices, Inc. [Electronic resource]. – Mode of access: <https://docs.xilinx.com/v/u/2020.1-English/ug909-vivado-partial-reconfiguration>. – Date of access: 10.10.2023.

HEALTH TESTS HARDWARE IMPLEMENTATION FOR ENTROPY SOURCES

Burko L., Kaiky M., Ivaniuk A.

Department of Informatics, Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: {burkoliana, kaikymykhailo}@gmail.com, ivaniuk@bsuir.by

The work discusses the process of synthesizing a microprogrammed machine for testing the suitability of physical sources of entropy in accordance with the NIST SP 800-90B standard. A comparison was made of the hardware costs and performance of the resulting microprogrammed machine and standard implementations of the selected tests. The developed firmware makes it possible to reduce hardware costs for hardware implementation of tests for true random number generators.

INTRODUCTION

The topic of random number generation is relevant in the modern world. There are millions of transactions every second that require encryption. To create true random number generator is a non-trivial problem. Most high-level programming languages use pseudo-random modules for number generation. For example they use system time as seed. Such modules can be susceptible to hacking or failures. To solve this problem there are various methods of testing random number sources in real time. For example NIST Special Publication 800-90B [1], BSI AIS20/31 compliant tests.

I. ENTROPY SOURCE STRUCTURE

In modern True Random Number Generators entropy sources are implemented as a set of physically unclonable functions that form N -channel sequences of random bits. Each channel requires a separate block of health tests for correct operation according to NIST 800-90B [1]. The entropy source model is represented in figure 1 in detail.

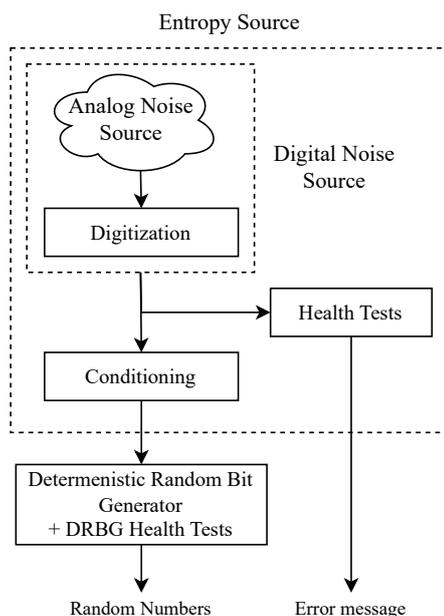


Figure 1 – Entropy Source model

The noise source is the root of security for the entropy source and for the Random Bit Generator as a whole. This is the component that contains the non-deterministic, entropy-providing process that is ultimately responsible for the uncertainty associated with the bitstrings output by the entropy source.

The optional conditioning component is a deterministic function responsible for reducing bias and/or increasing the entropy rate of the resulting output bits (if necessary to obtain a target value).

Health tests are an integral part of the entropy source design that are intended to ensure that the noise source and the entire entropy source continue to operate as expected.

II. HEALTH TESTS DESCRIPTION

There are 3 types of tests.

1. *Start-up health tests* are designed to be performed after powering up, or rebooting, and before the first use of the entropy source. They provide some assurance that the entropy source components are working as expected before they are used during normal operating conditions, and that nothing has failed since the last time that the start-up tests were run.

2. *Continuous health tests* are run indefinitely on the outputs of the noise source while the noise source is operating. Continuous tests focus on the noise source behavior and aim to detect failures as the noise source produces outputs. The purpose of continuous tests is to allow the entropy source to detect many kinds of failures in its underlying noise source.

3. *On-demand health tests* can be called at any time. NIST 80-900B [1] does not require performing testing during operation.

NIST 80-900B provides two approved health tests: *the Repetition Count test* and *the Adaptive Proportion test*. If these two health tests are included among the continuous health tests of the entropy source, no other tests are required.

The goal of *the Repetition Count Test* is to quickly detect catastrophic failures that cause the noise source to become "stuck" on a single output value for a long period of time. A variable C - considered critical.

$$C = 1 + \left\lceil \frac{-\log_2(a)}{H} \right\rceil \quad (1)$$

where H – the min-entropy of the samples from a (digitized) noise source or of the output from an entropy source, a – The probability of falsely rejecting the null hypothesis (type I error).

The *Adaptive Proportion test* checks the ratio of zeros and ones in each window.

Like all statistical tests, both of these tests have a false positive probability – the probability that a correctly functioning noise source will fail the test on a given output. In many applications, a reasonable choice for the probability of type I error is $a = 2^{-20}$.

For example, for $a = 2^{-20}$, an entropy source with $H = 2.0$ bits per sample would have a repetition count test cutoff value of $1 + 20/2 = 11$.

III. PRACTICAL RESULTS

Dataset, that comes on Health tests stand divided into blocks with size W . According to Health tests all blocks are checked separately and the validity result is issued within each block. The window size W is selected based on the alphabet size, and shall be assigned to 1024 if the noise source is binary and 512 if the noise source is not binary. In this work was decided to check also $W = 4096$. P will be used as the evaluation parameter, that shows amount of windows processing per second.

$$P = f_{CLK} \cdot W, \quad [\text{Giga Windows/s}] \quad (2)$$

where W – window size, f_{CLK} – maximum synchronization frequency.

Table 1 – Health Tests Hardware Utilization and Performance

Name	Resources				
	W	LUT	FF	f_{CLK} [MHz]	P
Adaptive	512	15	10	246	125
Repetition Count	512	21	13	239	122
Adaptive	1024	17	11	241	244
Repetition Count	1024	27	16	232	237
Adaptive	4096	21	14	238	974
Repetition Count	4096	34	30	230	934

IV. CONCLUSION

During the work, standards for constructing Health Tests for hardware true random number generators were studied, two single-channel hardware tests for the entropy source and a specialized stand for their validation on a Xilinx FPGA were designed. Prototyping stand is in figure 2. The results are presented in table 1. Dependences of hardware costs on the size of the working window for Health Tests were obtained. With an increase in the number of random bit channels, it is necessary to ensure that Health Tests operates in multi-channel mode, which entails an increase in hardware costs and power consumption of the chip. Further research involves studying methods for processing random sequences of multi-channel entropy sources to implement multi-channel Health Tests with reduced hardware costs and power consumption.

V. BIBLIOGRAPHY

1. NIST Special Publication, NIST SP 800-90B 3pd, Recommendation for the Entropy Sources Used for Random Bit Generation – NIST, January 2018. [Electronic resource] – Mode of access: <https://csrc.nist.gov/external/nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90B.pdf>. – Date of access: 15.10.2023.

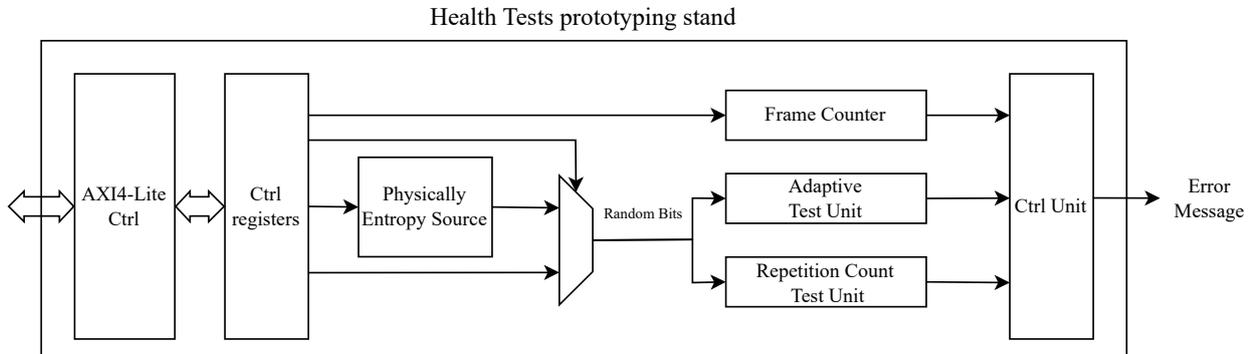


Figure 2 – Structure of Health Tests fast prototyping stand

SOFTWARE FOR REGISTERS STRUCTURE HDL-DESCRIPTION GENERATION WITH AMBA PROTOCOLS SUPPORT

Burko L., Kaiky M., Tushinskaya E.

Department of Informatics, Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus

E-mail: {burkoliana, kaikymykhailo}@gmail.com, tushkan@fksis.by

This work discusses a software tool for automatically building and checking registry structures with programmatic access using AMBA protocols. A specialized tool has been developed for constructing an HDL description of register structures and a verification environment with the generation of test sequences for testing them.

INTRODUCTION

In the modern system-on-chip (SoC) design stack, software-accessible registers for accessing peripheral devices and hardware accelerators play an important role. Modern SoCs have dozens of hardware accelerators in their composition, which are accessed using various interfaces. An example of such interfaces is the AMBA (Advanced Microcontroller Bus Architecture) family of protocols proposed by ARM and including the following interfaces: AXI3, AXI4, AHB, APB, ASB. Despite this wide variety of interfaces, the model of register memory that is accessed is unchanged and is a set of numbered cells. Due to the necessity of designing various register models for hardware accelerators, an approach to automatically generate such structures was proposed. In addition to generating register structures, the issue of verifying the resulting HDL descriptions using the SystemVerilog is also considered.

I. STRUCTURE OF SOFTWARE-ACCESSIBLE REGISTERS HARDWARE IMPLEMENTATION

In the general case, hardware implementations of software-accessible registers consist two fundamental components: an interface controller (for example, AXI4-Stream) and the memory itself. A memory block can be implemented in various ways, for example based on registers or SRAM cells. The implementation of a memory block depends on the requirements specified by the developers, such as read/write latency, the need to reset the value of cells after read/write operations, etc. In this paper, the authors propose a universal structure for describing an interface controller and a memory block for implementing a software-accessible unit. The proposed HDL-description consists of a reconfigurable interface controller, which allows, at the design stage, to select an implementable interface for communication with the host controller. The memory block is connected to the interface controller using a native interface (always remains unchanged), with the following signals: en, we, addr, wdata, rdata. Using this design approach, we can replace only the interface part without making significant adjustments to the memory structure.

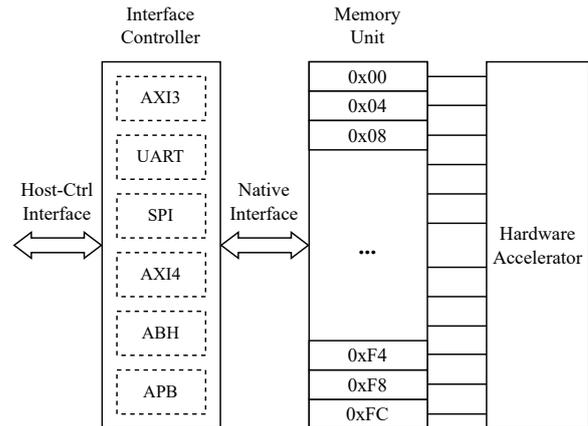


Figure 1 – Software-accessible Registers Structure

II. GENERATING MEMORY STRUCTURES

It was proposed to obtain an HDL description of memory structures using the developed generator. This generator operates based on a set of constraints for memory requirements, and the generator output is an HDL description in SystemVerilog for the memory unit. Constraints and requirements for a memory unit are specified in the following format:

```
'define RST_TYPE sync
'define RST_LEV 1
'define MEM_TYPE DFF
'define DFF_SYNC rising_edge
'define MEM_HOST_WLATENCY 2
'define MEM_HOST_RLATENCY 1
'define MEM_HW_WLATENCY 1
'define MEM_HW_RLATENCY 0
'define status // 32 // W/R // R // 32'h12
```

When such a description is transmitted to the generator, a memory will be generated based on D-flip-flops featuring synchronous high-level reset, write/read latency for the host controller - 2/1 system clock cycles, for a hardware accelerator - 1/0, and will contain one A 32-bit register named status. The access mode for the host controller is read and write, the access mode for the hardware is read only, the value at initialization/reset is 12 in hexadecimal.

III. VERIFICATION OF SOFTWARE-ACCESSIBLE REGISTERS STRUCTURE

To verify the generated structure, together with the interface controller, a series of test benches are recommended, simulating register access through a specified protocol (for example, AXI4-Lite or UART). For this purpose, "automatic tasks" were developed in the SystemVerilog language. Generation of addresses and data during read/write operations is carried out using root cases and the random function. Test sets are supplied to the block of software-accessible registers in such a way as to cover all possible modes of access to registers (read/write from the host controller and hardware accelerator) and cover all switching of the address decoder in the memory unit. After passing the verification stage, only when coverage of switching in the decoder and operating modes is achieved can a conclusion be made about the functional correctness of the generated structure.

An example of such a design for a write operation using the AXI4-Lite protocol (fig. 2):

```
task automatic axi_write;
input [S_AXI_ADDR_WIDTH - 1 : 0] addr;
input [S_AXI_DATA_WIDTH - 1 : 0] data;
begin
s_axi_wdata = data;
s_axi_waddr = addr;
s_axi_wvalid = 1;
s_axi_wstrb = 4'b1111;
wait(s_axi_awready && s_axi_wready);

@(posedge s_axi_aclk) #1;
s_axi_wstrb = 4'b0000;
s_axi_wvalid = 0;
s_axi_wvalid = 0;
end
endtask
```

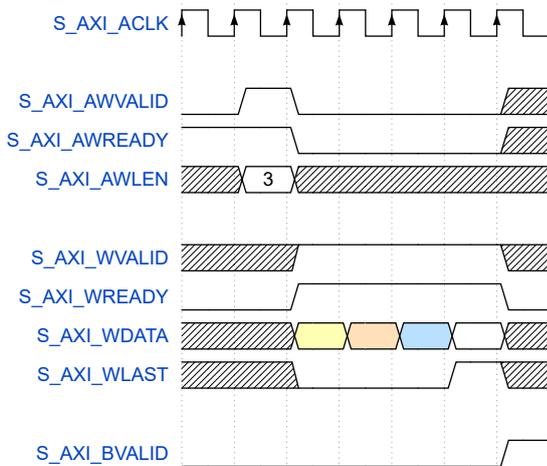


Figure 2 – AXI4-Lite Write Sequence

Read operation using the AXI4-Lite protocol (fig. 3):

```
task automatic enforce_axi_read;
input [S_AXI_ADDR_WIDTH - 1 : 0] addr;
input [S_AXI_DATA_WIDTH - 1 : 0] expected_data;
begin
s_axi_araddr = addr;
s_axi_arvalid = 1;
s_axi_rready = 1;
wait(s_axi_arready);
wait(s_axi_rvalid);
if (s_axi_rdata != expected_data) begin
display("Error: Mismatch in AXI4
read at %x:", addr, "expected %x,
received %x", expected_data, s_axi_rdata);
end
@(posedge s_axi_aclk) #1;
s_axi_arvalid = 0;
s_axi_rready = 0;
end
endtask
```

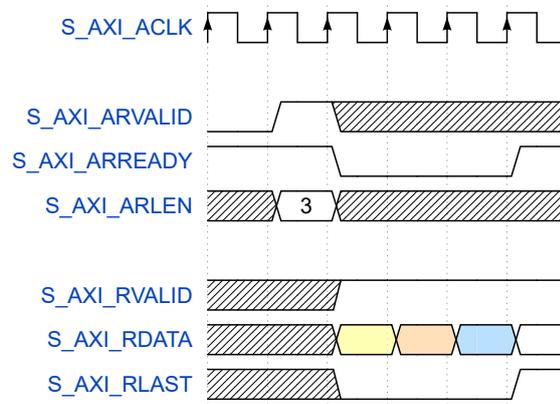


Figure 3 – AXI4-Lite Read Sequence

IV. CONCLUSION

A generator of HDL descriptions was developed for the structures of software-accessible registers with support for an arbitrary memory unit structure and various interface controllers (UART, SPI, AXI4-Lite). The generator receives as input a file with restrictions and parameters of the register memory model, and generates an HDL description of the implementation of a memory unit and a set of test sequences in the SystemVerilog.

1. Advanced Microcontroller Bus Architecture (AMBA), ARM Computing [Electronic resource] – Mode of access: <https://developer.arm.com/Architectures/AMBA>. – Date of access: 11.10.2023.
2. Design and Performance Analysis of 32 × 32 Memory Array SRAM for Low-Power Applications, Xue, Xingsi and Sai Kumar, Department of Electrical Engineering, Linköping University, 2022 [Electronic resource] – Mode of access: https://www.researchgate.net/publication/368381578_Design_and_Performance_Analysis_of_32_32_Memory_Array_SRAM_for_Low-Power_Applications. – Date of access: 15.10.2023.

RANDOM NUMBER GENERATION ON RECONFIGURABLE RING OSCILLATOR

Kaiky M., Ivaniuk A.

Department of Informatics, Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: kaikykhailo@gmail.com, ivaniuk@bsuir.by

This work is devoted to the study of a reconfigurable ring generator based on two-input XOR elements used to generate true random numbers. The paper examines a system for recording generator frequency oscillations and patterns of changes in oscillation frequencies at its outputs depending on the given generator configuration and the temperature of the FPGA chip. An approach to generating truly random numbers using a counter of the leading edges of the generated signal is considered.

INTRODUCTION

According to the NIST SP 800-90 [1] group of standards, sources of truly random numbers must be built on the basis of a physical noise generator, followed by digitization and processing of the random signal, converting it into sequences of random numbers. The most commonly used source of such noise is usually inverter-based ring oscillators.

I. STRUCTURE OF RECONFIGURABLE RING OSCILLATOR

The work [2] proposes the structure of a ring oscillator based on N two-input logical XOR gates (fig. 1). Inputs C are generator configuration inputs, EN is a control input, and when its level is active, the generator can begin to oscillate (depending on the configuration). The proposed generator design begins to oscillate only when the Hamming weight of the vector C is odd, and the oscillation frequency depends on the number of units and their position in the configuration. The gates were implemented on the FPGA element base of Xilinx (Zynq-7000 family) and are LUT2 elements.

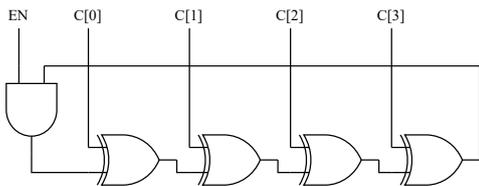


Figure 1 – Reconfigurable ring oscillator

II. MEASURING OSCILLATIONS FREQUENCIES

The oscillator frequencies are measured using two counters - a window counter tmw_{cnt} , a counter of the leading edges of generator oscillations clk_{cnt} . For the developed generator, the frequencies of its operation were measured on various configurations (fig. 2). Each configuration was measured 1000 times to obtain the average oscillation frequency. Measurements were carried out on a ring oscillator with 8 XOR elements, corresponding to 128 different configurations and on a window of 2 microseconds ($tmw_{cnt} = 100000$ at a system frequency of 50 MHz).

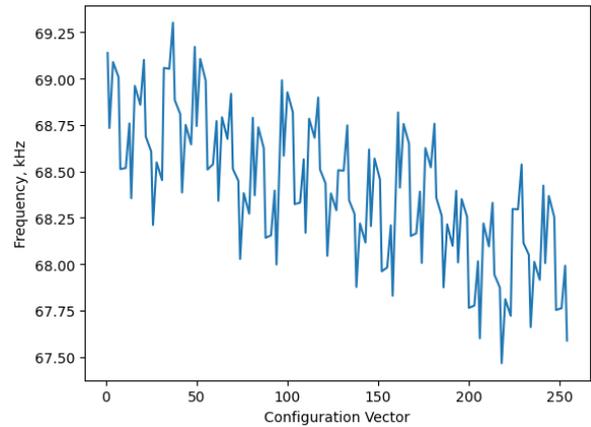


Figure 2 – Generator oscillation mean frequency depending on configuration

As can be seen from the figure, the frequency dependence tends to decrease the oscillation frequency with increasing configuration value. This is due to the number of toggled XOR elements in the generator. For greater clarity, we present the dependence of oscillation frequencies on the count of "ones" in the configuration vector (table 1). Also, clear patterns are visible, they are associated with combinations of switchable elements XOR.

Table 1 – Dependence of oscillation frequencies

F, [kHz]	Count of ones in configuration			
	1	3	5	7
f_{min}	68.506	67.764	67.466	67.588
f_{mean}	68.848	68.555	68.253	67.953
f_{max}	69.138	69.301	68.989	68.290

III. REASONING FOR CHOOSING THE SIZE OF THE MEASUREMENT TIME WINDOW

In the experiment described above, the window size was 2 microseconds; this value was not chosen by chance. In the experiment described above, to determine the window size, generator was run 10 times with windows of sizes: $10^0, 10^1, 10^2, \dots, 10^9$ toggle the tmw_{cnt} . As a result, we obtain the following dependence of the number of decimal places on the window size: 1, 1, 2, 2, 4, 5, 6, 7, 8, 9.

IV. FREQUENCY COUNTER AS A RANDOM NUMBER SOURCE

In addition to its main purpose, the oscillation frequency counter can also be used as a post-processing circuit for random sequences. To investigate this possibility, each configuration was fed to the generator 10,000 times. The figure 3 shows the distribution of numbers for 8 of 128 generator configurations, as you can see – they are shifted relative to each other and have different shapes.

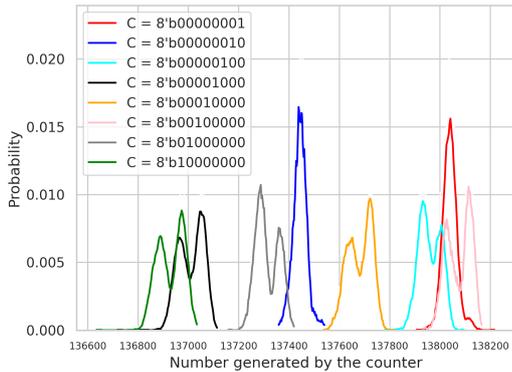


Figure 3 – Probability of numbers generated by the counter

Bias is the result of the special behavior of the counter as a post-processing circuit. Since the size of the measurement window is a fixed number, some of the higher digits of the counter will not change their value, while the lower ones, on the contrary, will make noise. This hypothesis is confirmed by calculating the probability of the appearance of one and zero in the digits of this counter (fig. 4).

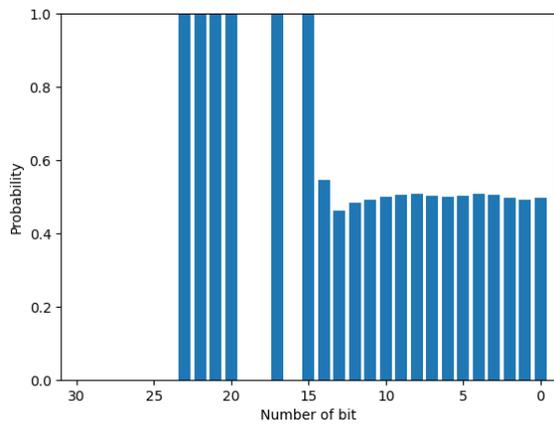


Figure 4 – Probability ones on clock counter

As can be seen from the figure 4, the highest 15 digits are stable and do not change their value depending on the experiment. The high stability of the most significant digits determines the generation of numbers in a given range (fig. 4). Introduce the value K - the number of unstable (oscillating) bits in the vector of frequency counter values. Then $K = 15$. Minor 15 - strives for the probability of occurrence of "ones" to be 0.5, which allows you to use data bits to generate truly random sequences.

V. RING OSCILLATOR TEMPERATURE MEASUREMENT

As world practice has shown, the oscillation frequency of a ring oscillator strongly and linearly depends on the temperature of the chip. To conduct an experiment on measuring the oscillation frequencies of the proposed reconfigurable generator, its configuration was fixed $tmw_{cnt} = 100000$, $C = 8'h1$. The temperature was changed using a temperature chamber (TestEquity Model 155) in the range from -15 to +115 degrees Celsius. The dependence of the generator oscillation frequency on temperature is shown in the figure 5. As can be seen from the figure, when the operating temperature range of the microcircuit is exceeded, the K value undergoes a change.

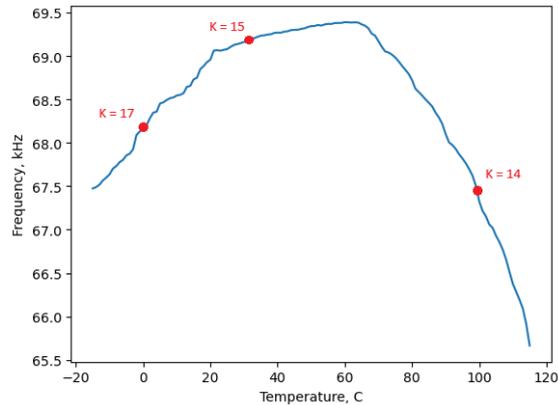


Figure 5 – Generator oscillation frequency depending on temperature

VI. CONCLUSION

When considering the structure of a reconfigurable ring oscillator, it is possible to limit the frequency-controlled oscillations and regulate their circuits by the configuration supplied to the operating inputs of the oscillator. The dependences of the generator oscillation frequency on a given configuration and temperature have been studied. The possibility of using a rising edge counter as a post-processing circuit for random bits is considered. It has been proven that the low-order bits of a frequency counter tend to have a uniform distribution of zeros and ones, which allows them to be used as true random number generators. The behavior of the edge counter in conjunction with a ring oscillator changes depending on the temperature of the microcircuit.

1. NIST Special Publication, NIST SP 800-90, Recommendation for Random Bit Generation – NIST. [Electronic resource] – Mode of access: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-90r.pdf>. – Date of access: 01.10.2023.
2. Yarmolik V.N., Ivaniuk A.A., Shynkevich N.N. Physically unclonable functions with controlled propagation delay. Informatics. 2022;19(1):32-49. (In Russ.) <https://doi.org/10.37661/1816-0301-2021-19-1-32-49>

FAST PROTOTYPING OF RECONFIGURABLE TRUE RANDOM NUMBER GENERATION IP-CORE

Kaiky M., Shamyna A., Ivaniuk A.

Department of Informatics, Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: kaikymykhailo@gmail.com, {shamyna, ivaniuk}@bsuir.by

This work is devoted of the structure of a reconfigurable IP-core for rapid prototyping of true random number sources based on Xilinx FPGAs. The developed IP-core allows you to configure TRNG at the design stage. In the proposed IP-core, developers can easily change the structure of the Entropy Source unit, Conditioning units, post-processors, performance tests, DRBG, component access interfaces.

INTRODUCTION

There are two types of random number generators: pseudo-random and true random. A pseudorandom number generator is an algorithm that simulates a random number. Predicted sequences can be generated by the pseudo-random number generator in software implementations, but it cannot be used as an entropy source for cryptographic methods. A true random number generator uses physical processes (eg. thermodynamic noise, mechanical noise, quantum process noise) to generate random numbers that cannot be predicted. These generators are more expensive and more difficult to use, but they produce more cryptographically secure random numbers.

I. NIST STANDARDS FOR THE CONSTRUCTION OF RANDOM NUMBER GENERATORS

NIST Special Publications (SP) 800-90 series describe methods and approaches for constructing random number generators for cryptographic and non-cryptographic purposes. SP 800-90A [1] defines several characteristics of a Deterministic Random Bit Generator (DRBG) based on cryptographic algorithms, SP 800-90B[2] contains recommendations for the development and testing of entropy sources, SP 800-90C [3] – identify Random Bit Generator (RBG) implementation designs that incorporate DRBG mechanisms as specified in SP 800-90A and use an entropy source as specified in SP 800-90B.

According to NIST SP 800-90C, true random number generators can be of three types:

1. RBG1: The entropy source and DRBG are implemented in different blocks and are connected to each other using a secure channel.
2. RBG2: Entropy source, random bit post-processing block and DRBG are within the same safety boundary.
3. RBG3 (XOR): The design contains one or more verified entropy sources and DRBGs whose the output is XORed to produce random numbers with full entropy.

Despite the different ways of constructing generators, they all have a common structure (fig. 1), consisting of a physical source of entropy, health tests of the source, a conditioning unit and DRBG.

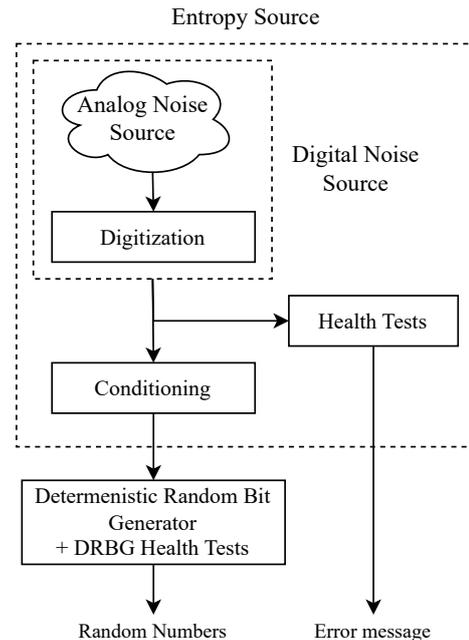


Figure 1 – NIST Random Bit Generator Structure

II. DEVELOPMENT A RECONFIGURABLE RANDOM NUMBER GENERATOR

Accelerate the process of prototyping and researching sources of true random numbers, an approach was proposed to develop a reconfigurable generator that allows replacing blocks of the physical source of entropy, conditioning unit, DRBG, changing the nature of the connections between them (both in accordance with NIST SP 800-90C [3], and in other methods). The general structure of the developed IP-core of such a reconfigurable generator is depicted in figure 2, blocks circled in a dotted line – support reconfiguration at the design stage (before synthesis). The developed generator, as in NIST SP 800-90C [3], consists of the following main parts: entropy source, conditioner, Health Tests Unit, DRBG, but there are also other blocks, such as – hardware FIFO for storing and speeding up access to random numbers, interface controller blocks and registers programmatically accessible to the host controller. The developed number generator can be integrated into a system on a chip, such as an FPGA or ASIC,

and provide host-controller (processor) access to the source of random numbers using various interfaces. The interface is configured at the design stage using the `define` directives of the SystemVerilog hardware description language and currently the following interfaces are supported: AXI4-Lite, Naive Interface (addr, data, we, re), UART, SPI, and others can also be added.

The main goal in the development of TRNG is to make the right choice of the physical source of entropy and to ensure the reliability of random number generation using post-processing methods of data from this source. Usually, the construction of ring generators is used as a source of entropy [4], linear feedback registers (LFSR) are used as a Conditioner, and NIST AES256/AES512 algorithms are used as DRBG.

III. CONCLUSION

An IP-core of a hardware accelerator based on Xilinx FPGA was developed to conduct experimental studies of cryptographic primitives, in particular, physically unclonable functions as sources of entropy, post-processing blocks of random data and deterministic random number generators (DRBG), which underlie the hardware implementation of the generator true random numbers according to NIST standards. The oscillation frequency of ring generators strongly and linearly depends on the temperature [5], which makes it difficult to use them as sources of

entropy, reduces reliability and resistance to attacks on TRNG. The developed IP-core makes it possible to introduce various entropy source structures into the TRNG structure, which makes it possible to expand the study area and conduct research on other physical sources, such as static memory (SRAM).

BIBLIOGRAPHY

1. NIST Special Publication, NIST SP 800-90A 3pd, Recommendation for Random Bit Generator (RBG) Constructions – NIST, September 2022. [Electronic resource] – Mode of access: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90Ar1.pdf>. – Date of access: 15.10.2023
2. NIST Special Publication, NIST SP 800-90B 3pd, Recommendation for Random Bit Generator (RBG) Constructions – NIST, September 2022. [Electronic resource] – Mode of access: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90B.pdf>. – Date of access: 15.10.2023
3. NIST Special Publication, NIST SP 800-90C 3pd, Recommendation for Random Bit Generator (RBG) Constructions – NIST, September 2022. [Electronic resource] – Mode of access: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90C.3pd.pdf>. – Date of access: 15.10.2023.
4. Physically unclonable functions based on a controlled ring oscillator, Ivaniuk.A, Yarmolik.V
5. Temperature Compensation in CMOS Ring Oscillator, Dingyufei Zhang, Xiaohua Wei, Department of Electrical Engineering, Linköping University, 2022 [Electronic resource] – Mode of access: <https://liu.diva-portal.org/smash/get/diva2:1662148/FULLTEXT01.pdf>. – Date of access: 15.10.2023.

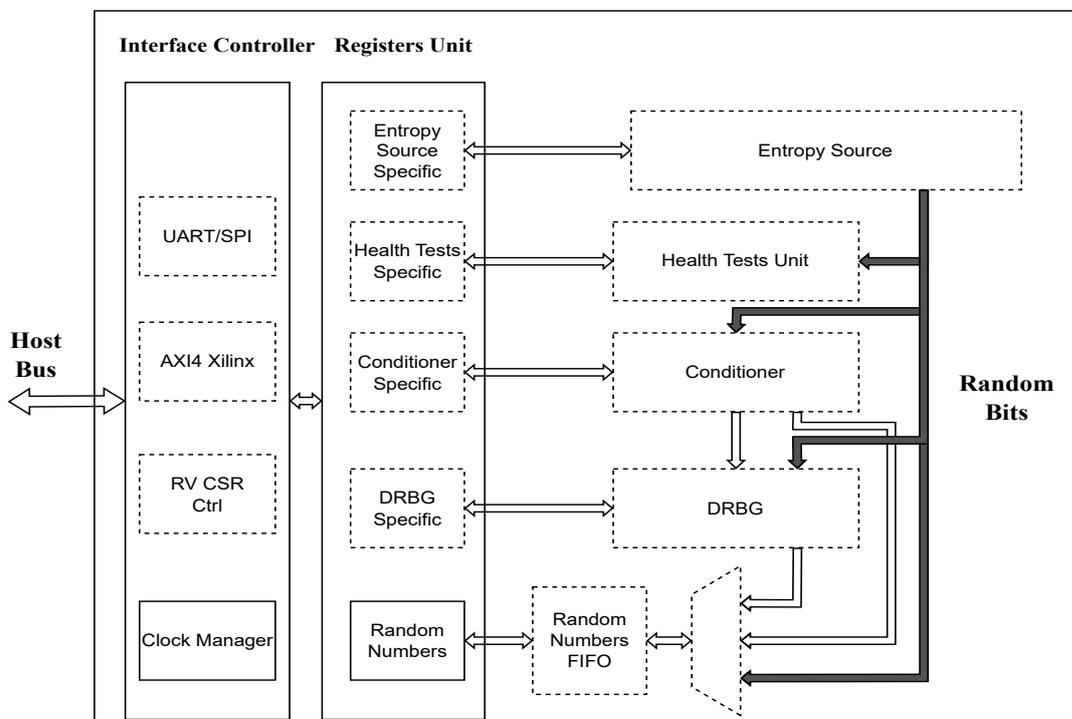


Figure 2 – Block diagram of the IP-core the reconfigurable true random number generator

RISC-V HARDWARE MODIFICATION FOR M-SEQUENCES GENERATION

Petrovsky D., Ivaniuk A.

Faculty of Computer Systems and Networks, Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus

E-mail: petrovsky.dmitr@gmail.com, ivaniuk@bsuir.by

A hardware modification of the soft processor core of the open RISC-V architecture to accelerate the generation of M-sequences is being considered. The results of a comparative analysis of the performance of completely software algorithms and an algorithm with support for hardware modification are shown, and the hardware costs for implementation in a Xilinx-7 FPGA chip are calculated.

INTRODUCTION

The main advantage of processor systems is the ability to perform a wide range of tasks, such as various mathematical algorithms and control systems. To speed up the execution of which, the following approaches can be used: software optimization, adding specialized co-processors, and a combination of hardware and software modifications. This article will discuss the third approach.

I. PROBLEM STATEMENT

As a processor core we will use the currently popular RISC-V[1] architecture, into the structure of which hardware changes will be introduced. This architecture supports instruction sets with different bit depths as shown in Table 1.

Table 1 – Basic instruction sets

Abbreviation	Name
RVWMO	Basic memory consistency model
RV32I	Basic set with integer operations, 32-bit
RV64I	Basic set with integer operations, 64-bit
RV128I	Basic set with integer operations, 128-bit

In addition to the basic sets, this architecture also contains additional sets that expand the functionality, as shown in Table 2.

Table 2 – Standard unprivileged command sets

Abbreviation	Name
M	Integer Multiplication and Division
F	Single-Precision Floating-Point
D	Double-Precision Floating-Point
V	Vector Operations

A pseudo-random sequence (PRS) is a sequence of numbers that was calculated according to some arithmetic rule, but has all the properties of a random sequence. One of the devices for generating PRS is a linear feedback shift register (LFSR), in which the value of the input (pushed) bit is equal to a linear Boolean function of the values of the remaining bits of the register before the shift. There are two types of feedback in LFSR, external feedback (Fibonacci configuration Figure 1.a) and internal feedback (Galois configuration Figure 1.b). Both

implementations have the same generator functions, but the Galois configuration, due to the ability to parallelize XOR operations, allows for greater performance in both software and hardware implementations.

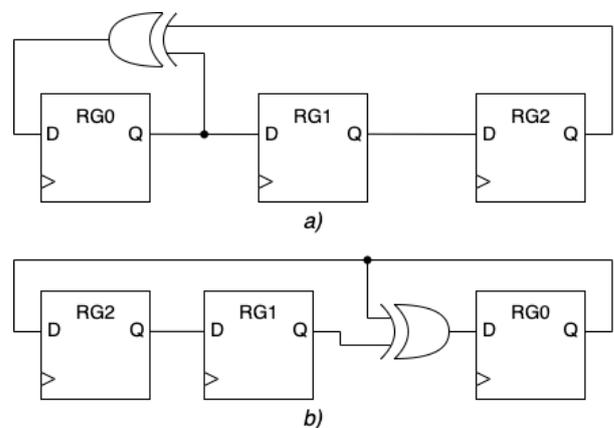


Figure 1 – Shift register with linear feedback a) Fibonacci configuration b) Galois configuration

The PRS with the maximum length is called an M-sequence and has the following properties:

- M-sequences are periodic with period $N = 2^n - 1$;
- the number of symbols taking the value one, over the length of one period of the M-sequence, is one more than the number of symbols taking the value zero;
- any combination of symbols of length n in one period of the M-sequence, with the exception of a combination of n zeros, occurs no more than once;
- a combination of n zeros is prohibited: only a sequence of only zeros can be generated on its basis.

The extensive use of M-sequences in cryptography problems, communication systems and digital signal processing algorithms makes it urgent to solve issues of optimization and acceleration of PRS generation algorithms.

II. RESEARCH RESULTS

A pipeline implementation of a soft processor of the RISC-V architecture with support for the

RV32I[2] instruction set was taken. This implementation contains a set of 32 registers and includes 39 instructions: fetch-store, logical and arithmetic operations, conditional and unconditional jump instructions.

To speed up M-sequence calculations, it is proposed to add additional logic to the general-purpose register and modify the command decoding device.

The additional logic for the general-purpose register will allow it to perform both the LFSR and general-purpose register functions. To make the general-purpose register work as a shift register between triggers we will place a two-input multiplexer, the selecting signal of which provides a choice between parallel write or shift operation. In the resulting shift register input bit is directed from the stage of elements AND/XOR allowing the implementation of external feedback in the LFSR. This structure(Figure 2) allows you to implement any polynomial to form a PSP with a dimension of up to 32-bit. The polynomial is specified in the 32-bit status control register (SCR), so a one set in bit number N adds the N+1 term to the polynomial. For example, a polynomial of the form « $x^{32} + x^{22} + x^2 + x^1 + 1$ » will correspond to the value of register 0x8020003. If this register takes a non-zero value, then the general-purpose register switches its mode to LFSR, if this SCR register took a zero value, the register functions as a normal general-purpose register.

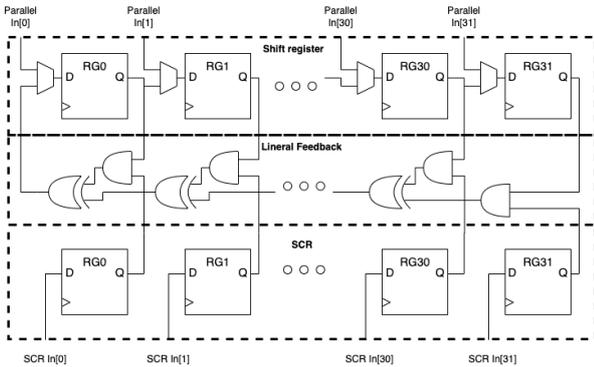


Figure 2 – LFSR block structure

Modification of the control device is an additional reading signal. This signal is generated for the register we have selected while executing commands that use this register.

This hardware modification allows the register with LFSR function to be used for operation with standard processor instructions. Thus, all logical operations, arithmetic operations and memory write operations will cause the next LFSR value to be calculated, so that a new value can be obtained every processor clock.

Here are the results of comparison of software implementations with our hardware modification. The Prog.1 fully corresponds to our hardware modification, it checks all bits of the polynomial in a loop

and when the value of the bit "1" to perform the XOR operation, the shift always occurs. The Prog.2 computes only one polynomial, runs without a loop and without unnecessary shift operations. Table 3 shows for comparison the number of CPU cycles required to calculate one element of the M-sequence of the following polynomials «0x8020003, 0x8579D037, 0xFFFFFFFFEE». The general graph of clock cycles consumed for all polynomial lengths is shown in Figure 3.

Table 3 – Comparison CPU cycles

Polynomial	Number of terms	Prog.1	Prog.2	Prog.3 with mod.
0x8020003	4	312	12	1
0x8579D037	16	309	36	1
0xFFFFFFFFEE	30	295	64	1

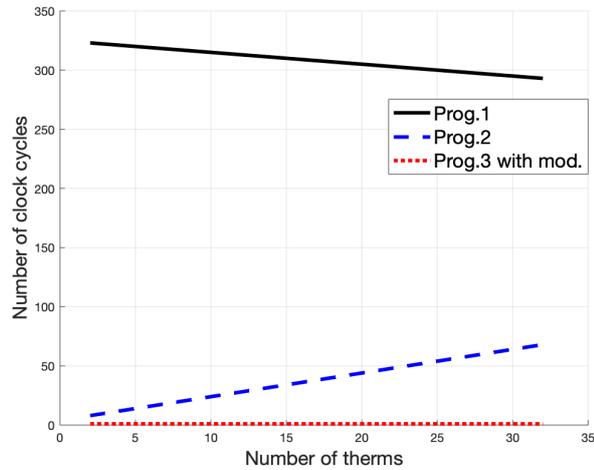


Figure 3 – Graph of the clock cycles

The RISC-V soft processor was implemented in a Xilinx-7 series FPGA chip, Table 4 shows the hardware overheads of the implementation without and with hardware modification, and the difference in their expenses.

Table 4 – table

Implementation	Slice	LUT's	Registers
Original RISC-V	704	1489	1433
Mod. RISC-V	789	1638	1465
Difference	85	149	32

The data in Tables 3-4 show that fully software implementations are inferior to hardware-software implementations in performance by a factor of 8 to 320. Hardware costs are not significant compared to the resources occupied by the soft processor core.

1. Official site RISC-V International [Electronic resource] – Mode of access: <https://riscv.org>. – Date of access: 26.10.2023.
2. Harris. S. L. Digital Design and Computer Architecture, RISC-V / S. L. Harris, D. Harris; – Morgan Kaufmann. Press, 2021. – 592 p

Секция «Математическое
моделирование и компьютерная
графика»

Section «Mathematical modeling and
computer graphics»

АЛГОРИТМ УПРАВЛЕНИЯ ТРАЕКТОРИЕЙ ДВИЖЕНИЯ ИГРОВОГО ПЕРСОНАЖА

Баштык П. А., Кукин Д. П., Шабанович Р. А.

Кафедра вычислительных методов и программирования, кафедра систем управления,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь

E-mail: p.bashtyk@bsuir.by, kudin@bsuir.by, r.shabanovich@bsuir.by

В этой статье авторы предлагают новый алгоритм управления траекторией движения игрового персонажа, основанный на различных методах оптимизации. Алгоритм позволяет генерировать реалистичные и эффективные траектории движения в разнообразных игровых сценариях.

ВВЕДЕНИЕ

В современных компьютерных играх траектория движения игрового персонажа является важным фактором, влияющим на реализм, интерактивность и удовлетворенность игроков. Существует множество алгоритмов управления траекторией движения, которые используют различные подходы, такие как планирование, оптимизация, обучение с подкреплением и т.д. Однако, большинство из них имеют свои недостатки, такие как высокая вычислительная сложность, низкая гибкость, ограниченная обобщающая способность и т.д.

Целью данного исследования является разработка нового алгоритма управления траекторией движения игрового персонажа, который обладает высокой эффективностью, реалистичностью и адаптивностью. Для этого мы предлагаем комбинировать различные методы оптимизации.

I. АНАЛИЗ СУЩЕСТВУЮЩИХ АЛГОРИТМОВ

Анализ существующих алгоритмов управления траекторией движения игрового персонажа является важным шагом при разработке нового алгоритма. Рассмотрим некоторые из распространенных подходов и их особенности [1].

Путевые точки (Waypoints): этот подход основан на использовании заданных точек на карте, через которые персонаж должен пройти. Персонаж движется от одной точки к другой, следуя определенному пути. Преимущество такого подхода заключается в простоте реализации и предсказуемом поведении персонажа. Однако он может ограничивать свободу движения и не учитывать динамические изменения в окружающей среде.

Кинематические алгоритмы моделируют движение персонажа с использованием математических уравнений, учитывая его физические ограничения. Это позволяет достичь более реалистичного и плавного движения. Кинематические алгоритмы могут быть основаны на законах физики или моделировании анимации персонажа. Однако они могут быть сложны в реализации и требовать дополнительных вычислительных ресурсов.

Методы машинного обучения, такие как нейронные сети или генетические алгоритмы, могут использоваться для обучения персонажа на основе опыта или генерации оптимальной траектории. Это позволяет создавать более интеллектуальные и адаптивные алгоритмы управления. Однако использование методов машинного обучения может потребовать больших объемов данных и вычислительных ресурсов для обучения модели [2].

Навигационные сетки (Navigation Grids): этот подход основан на разбиении игрового пространства на сетку клеток, каждая из которых имеет определенное значение проходимости. Этот подход предоставляет более гибкое управление траекторией и учитывает препятствия в окружающей среде. Однако он может быть вычислительно сложным и требовать больших вычислительных ресурсов при больших игровых пространствах.

Алгоритм A-star основан на поиске оптимального пути от начальной точки к целевой точке в графе состояний игрового пространства [3]. Он использует комбинацию двух функций для оценки стоимости пути: функция $g(n)$, которая представляет стоимость пути от начальной точки до текущей точки, и функция $h(n)$, которая представляет эвристическую оценку стоимости пути от текущей точки до целевой точки. Сумма этих двух функций, $f(n) = g(n) + h(n)$, используется для выбора следующей точки в пути.

Преимущества алгоритма A* включают его эффективность и оптимальность при правильно выбранных эвристических функциях. Он может учитывать препятствия и динамические изменения в окружающей среде, а также обеспечивать гладкое и естественное движение персонажа.

Тем не менее, алгоритм A* имеет свои ограничения. Он может быть вычислительно сложным при работе с большими игровыми пространствами или сложными топологиями. Также выбор подходящих эвристических функций может быть сложной задачей, и неправильный выбор может привести к неоптимальным или нереалистичным путям [4].

В зависимости от конкретных требований игры и поведения персонажа, можно выбрать различные алгоритмы управления траекторией

движения или комбинировать их для достижения желаемого результата.

II. ПРЕДЛАГАЕМЫЙ АЛГОРИТМ

Предлагаемый алгоритм управления траекторией движения основан на комбинации алгоритма A^* и кинематического управления. Он позволяет персонажу эффективно и плавно перемещаться по игровому пространству, учитывая препятствия и динамические изменения в окружающей среде.

Основные компоненты и принципы работы предлагаемого алгоритма.

Генерация пути: используя алгоритм A^* , генерируется оптимальный путь от текущей позиции персонажа до целевой точки на игровой карте. Алгоритм A^* учитывает стоимость перемещения между клетками с учетом препятствий и эвристической оценки расстояния до цели. Это позволяет найти наиболее короткий и безопасный путь.

Кинематическое управление: после генерации пути используется кинематическое управление для обеспечения плавного движения персонажа по этому пути. Кинематическое управление моделирует движение персонажа, учитывая его физические ограничения, такие как максимальная скорость и ускорение. Оно генерирует управляющие сигналы, такие как угол поворота или скорость, чтобы персонаж следовал по заданному пути.

Обновление пути: во время движения персонажа могут возникать динамические изменения в окружающей среде, такие как появление новых препятствий или изменение целевой точки. В таких случаях алгоритм A^* периодически обновляет путь, чтобы учесть новые условия. Это позволяет персонажу адаптироваться к изменениям и продолжать двигаться по оптимальному пути.

Результаты экспериментов с предлагаемым алгоритмом управления траекторией движения могут варьироваться в зависимости от конкретной реализации и игрового сценария. Однако ожидается, что алгоритм будет обеспечивать следующие преимущества:

Оптимальность пути: алгоритм A^* позволяет находить оптимальный путь от текущей позиции до цели, учитывая препятствия. Это позволяет персонажу находить наиболее короткий и безопасный путь в игровом пространстве.

Плавное движение: кинематическое управление обеспечивает плавное и естественное движение персонажа по заданному пути. Это позволяет избежать рывков и неестественных перемещений, что способствует лучшему игровому опыту.

Адаптивность к изменениям: при возникновении динамических изменений в окружающей

среде, таких как новые препятствия или изменение целевой точки, алгоритм обновляет путь, чтобы учесть новые условия. Это позволяет персонажу адаптироваться к изменениям и продолжать двигаться по оптимальному пути.

Реалистичное движение персонажа: кинематическое управление учитывает физические ограничения персонажа, такие как максимальная скорость и ускорение, что позволяет достичь более реалистичного движения в игровой среде.

Конкретные результаты экспериментов с данным алгоритмом могут зависеть от множества факторов, включая сложность игровой сцены, точность моделирования физики, эффективность алгоритма и другие. Проведение реальных экспериментов и тестирование алгоритма в контексте конкретной игры может дать более точные и конкретные результаты относительно его производительности и эффективности.

III. ВЫВОДЫ

В данной статье представлен новый алгоритм управления траекторией движения игрового персонажа, который использует различные методы оптимизации. Основные достоинства алгоритма заключаются в следующем:

- Он позволяет генерировать реалистичные и эффективные траектории движения в разнообразных игровых сценариях, учитывая препятствия и динамические изменения в окружающей среде.
- Он обеспечивает плавное и адаптивное движение персонажа, используя кинематическое управление и периодическое обновление пути.
- Он основан на едином фреймворке, который интегрирует алгоритм A^* и методы обучения с подкреплением, что упрощает реализацию и настройку алгоритма.

Алгоритм может зависеть от множества факторов, таких как сложность игровой сцены, точность моделирования физики, эффективность алгоритма и другие. Поэтому, следует проводить дальнейшие исследования и тестирование алгоритма в контексте конкретной игры, чтобы получить более точные и конкретные результаты.

1. LaValle, S. M. (2006). Planning algorithms. Cambridge University Press.
2. Russel, S., Norvig, P. (2016). Artificial intelligence: a modern approach. Pearson.
3. Karaman, S., Frazzoli, E. (2011). Sampling-based algorithms for optimal motion planning. International Journal of Robotics Research, 30(7), 846-894.
4. Liao, T.-W., Shen, S. (2012). Fast motion planning for robot navigation. IEEE Transactions on Automation Science and Engineering, 9(4), 785-792.

ОСОБЕННОСТИ ПРИМЕНЕНИЯ ГЕОМЕТРИЧЕСКОГО МОДЕЛИРОВАНИЯ, НА БАЗЕ ГЕОМЕТРИЧЕСКИХ НОД, ПРИ РЕШЕНИИ ЗАДАЧ ПОСТРОЕНИЯ ТРЕХМЕРНЫХ МОДЕЛЕЙ

Горбатов Г. В., Воцьева Э. И.

Кафедра инструментального и прикладного программного обеспечения, МИРЭА – Российский технологический университет

Институт информационных технологий

Москва, Российская Федерация

E-mail: gorbatovg@mail.ru, eliboc@inbox.ru

В статье рассмотрены особенности применения геометрического (нодового) моделирования при построении классических трехмерных моделей, описаны основные подходы в реализации логических геометрических соединений, позволяющие создавать трехмерные объекты. Освещены плюсы и минусы данного подхода, а также использование геометрического моделирования в решении задач компьютерной графики.

ВВЕДЕНИЕ

Версия 2.92 2021 года вводит в открытое программное обеспечение Blender систему геометрического моделирования на основе узлов, предназначенных для манипуляции геометрией трехмерных объектов, называемая «Geometry Nodes». Данная технология основана на процедурных операциях и представляет возможность детальной параметризации создаваемой геометрии, что ускоряет и упрощает процесс моделирования, открывает новые возможности применения трехмерной графики для эффективного решения прикладных задач.

Целью данной публикации является исследование возможностей, особенностей и сфер применения нодовой системы для определения целесообразности использования машинного обучения как способа снижения порога вхождения в геометрическое моделирование.

I. ОСНОВНАЯ ЧАСТЬ

Структура геометрических нод имеет следующую логику: на вход и выход подаются определенные значения, различаемые посредством цветовой дифференциации. Путем комбинирования узлов и связей между ними создается трехмерная модель. Входы расположены в нижней левой части узла и предоставляют данные, необходимые узлу для выполнения своей функции. Каждый входной порт при отключении принимает значение по умолчанию, которое можно редактировать с помощью цветового, числового или векторного значения. Выходы расположены в верхней правой части узла, и их выходные данные могут быть переданы на входы других узлов (см. рис. 1).

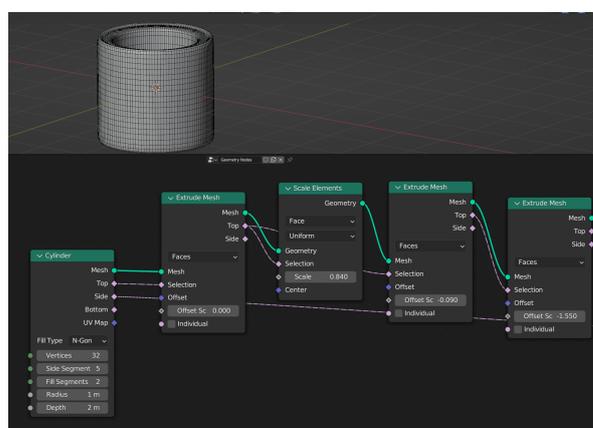


Рис. 1 – Пример модели, созданной с помощью геометрических нод

Некоторые узлы имеют специальные порты, которые могут принимать несколько входных данных. Эти порты будут иметь форму эллипса, а не круга, что указывает на их особое поведение. Порты в узлах различаются по используемым данным, каждое с уникальным цветом (см. рис. 2).

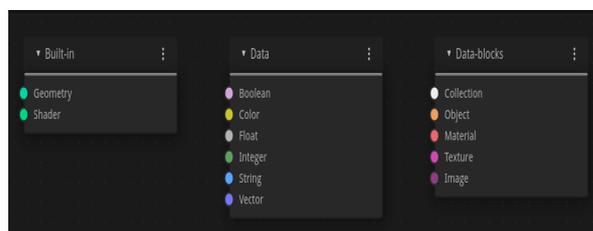


Рис. 2 – Цветовая дифференциация данных

Используемые данные входят в группу основных параметров, которые классифицируются тремя типами.

Основные наборы данных – данные о меше и шейдерах объекта, включают в себя:

- Geometry – геометрию объекта;
- Shader – шейдеры.

Типы данных – различные логические внутренние данные объекта, представляют собой

- Boolean – булевы значения;
- Color – цвета с наличием альфа-канала или без него;
- Float – числа с плавающей точкой. Данный порт также принимает значения (value map);
- Integer – целочисленные значения;
- String – текстовые значения;
- Vector – информация о векторе: направление, координаты и нормали.

Группы данных – второстепенные параметры, отвечающие за структуру проекта и внутренних объектов шейдера, состоят из:

- Collection – коллекций объектов;
- Object – объектов;
- Material – материалов;
- Texture – текстур;
- Image – изображений.

Помимо этого, к важным аспектам данных внутри любой трехмерной модели следует отнести параметры развертки – UV-координаты и анимацию по ключевым кадрам.

II. ОСОБЕННОСТИ И ПРОБЛЕМЫ НОДОВОГО МОДЕЛИРОВАНИЯ

Важно, что при использовании нодового подхода можно исключить ряд ошибок в топологии объектов, т.к. модель создается путем последовательного наложения логических и геометрических операций. Кроме этого, модели, спроектированные с помощью геометрии, имеют возможность динамической параметризации, т.к. вся их структура состоит из блоков данных, значения которых можно изменять в любое время, что недоступно в классическом моделировании и как следствие – возможность снижения трудозатрат на моделирование. Путем проектирования нодового дерева, выступающего в качестве “каркаса”, можно генерировать наборы моделей за короткое время [1], [2] через изменение параметров. Кроме того, стоит обратить внимание на важную особенность представления файлов в Blender: каждый проект может быть представлен в виде набора директорий. Наличие директории “Node Trees” позволяет многократно использовать нодовые деревья в других проектах и создавать библиотеки полученных геометрий. Также к числу достоинств геометрического подхода можно отнести оптимизацию машинных ресурсов. В ряде вариантов использования моделирование с применением нод превосходит традиционный метод по размеру итоговых файлов и затраченным расчетным мощностям [1].

Несмотря на перечисленные достоинства, освоение геометрического моделирования может представлять трудности для некоторых пользователей, в силу необходимости понимания типов

и структур данных и принципов представления и хранения объектов в Blender для проектирования нодовых деревьев. Помимо этого, стоит отметить и тот факт, что построение моделей путем создания нодовых связей возможно не для всех типов моделей, например таким способом будет сложно воссоздать модели, требующие особой точности, модели людей и ряд других, однако применение и возможная доработка такого подхода может позволить упростить и ускорить некоторые начальные этапы пайплайна в трехмерном моделировании.

III. Вывод

Исходя из проведенного анализа, можно выявить ряд проблем, возникающих при работе с геометрическим моделированием. В качестве решения проблемы освоения высокого порога вхождения в геометрическое моделирование можно предложить генерирование геометрии с помощью моделей машинного обучения, что позволит пропускать этап проектирования нодовых деревьев и сразу переходить к настройкам параметризации.

Создание нейронной сети, применяющей нодовое моделирование, может нести в себе ряд ключевых особенностей. Модели, генерируемые такой нейросетью, будут строиться по заранее подготовленным логическим параметрам – нодовой структуре моделирования, что несет в себе значительные преимущества по сравнению с существующими аналогами нейронных сетей для формирования трехмерных моделей. К таким преимуществам можно отнести следующие параметры:

- правильную топологию со сравнительно небольшим количеством полигонов, что значительно упростит дальнейшую работу с моделью;
- возможность доработки результирующих объектов благодаря наличию динамической параметризации;
- настроенную логическую составляющую объекта что позволит создавать аналогичные модели в любом проекте.

Текущей основной задачей для создания нейросети является сбор датасетов, на основе которых будет проводиться ее обучение.

СПИСОК ЛИТЕРАТУРЫ

1. Green, C. M. Optimizing game models: a comparison of geometry nodes and traditional modelling techniques in Blender. – 2023.
2. Gaillard, M. M Automatic differentiable procedural modeling / M. Gaillard, V. Krs, G. Gori, R. Méch, B. Benes // In Computer Graphics Forum. – 2022. –Vol. 41, №. 2. P. 289–307.
3. Blender 3.6 Manual [Electronic resource] / – Mode of access: <https://docs.blender.org/manual/en/latest/>. – Date of access: 17.10.2023.

РЕЙТИНГОВЫЕ СИСТЕМЫ И ИХ ВЛИЯНИЕ НА ПОДБОР ЦЕЛЕВОЙ АУДИТОРИИ ПРИ РАЗРАБОТКЕ КОМПЬЮТЕРНЫХ ИГР

Гуревич О. В., Кукин Д. П., Новицкая Л. И., Хмыз Д. Д.
Кафедра вычислительных методов и программирования,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: o.gurevich@bsuir.by

В статье разобраны основные рейтинговые системы, проводящие экспертизу и цензурирование игровых продуктов для европейских стран, США и России и рассмотрены требования, которые необходимо учитывать при разработке игр для определенных целевых аудиторий, разделенных по стандартным в геймдизайне возрастным группам.

ВВЕДЕНИЕ

При разработке компьютерных игр одним из ключевых элементов является проработка целевой аудитории. В индустрии компьютерных игр определение целевой аудитории начинается, как правило, еще до прототипирования, на этапе разработки игровой концепции. Довольно редко, но иногда бывают ситуации, когда определение целевой аудитории осуществляется уже после создания прототипа (обычно это характерно для инди-разработчиков или для компаний разработчиков на этапе пред-продакшена в случае вариативности графики/сеттинга или сюжета). В любом случае разработчикам игры необходимо учитывать требования рейтинговых систем и цензуры, которые в большинстве случаев основываются на принятых возрастных рейтингах.

I. РЕЙТИНГОВЫЕ СИСТЕМЫ

Разработчикам и издателям видеоигр нужно обращать внимание на требования законодательства о возрастных ограничениях в отношении игр. Чаще всего продажа и продвижение игр без рейтинга практически невозможны (особенно это касается крупных издателей и продавцов). Компании-производители консолей, цифровые магазины и ритейл-сети могут просто отказаться выпускать игру без маркировки, поскольку наличие рейтинга стало устоявшейся практикой в индустрии (хотя в большинстве стран мира, и, в частности, в СНГ, маркировка не является обязательным атрибутом).

Есть несколько основных возрастных рейтингов: американская ESRB, европейская PEGI, немецкая USK, японская CERO, российская RARS, а также еще несколько региональных организаций. Отличаются они регионами применения и особенностями работы.

В большинстве стран Европы действует система PEGI (Pan European Game Information) – европейская рейтинговая система компьютерных игр и другого развлекательного программного обеспечения. Была разработана Европейской федерацией интерактивного программного обеспе-

чения и начала работу в апреле 2003 года [1]. Для принятия и определения рейтингов для компьютерных видеоигр и другого развлекательного программного обеспечения в США и Канаде существует Entertainment Software Rating Board (ESRB) [2] – негосударственная организация, основное направление деятельности – принятие и определение рейтингов для компьютерных видеоигр и другого развлекательного программного обеспечения в США и Канаде. ESRB была основана в 1994 году.

Для России существует Возрастная классификация информационной продукции (Russian Age Rating System, RARS) представляющая собой совокупность норм, регулирующих доступ к информации, причиняющей вред здоровью и развитию детей в Российской Федерации. Введена с 1 сентября 2012 года, после вступления в силу федерального закона «О защите детей от информации, причиняющей вред их здоровью и развитию» [3].

Если игра планируется для выпуска только через цифровые магазины, то можно использовать рейтинговую систему IARC (International Age Rating Coalition). При создании страницы в цифровом магазине разработчику необходимо будет заполнить анкету с закрытыми вопросами о наличии «чувствительного» контента. На основе внутренних алгоритмов система присвоит игре рейтинг одновременно для большинства региональных рейтингов. IARC поддерживают Google Play, PlayStation Store, Microsoft Store, Nintendo eShop, Oculus, Origin. В отличие от рейтингов ESRB или PEGI, получение рейтинга IARC не является платной опцией. А если, по мнению разработчика, рейтинг IARC неправильный, он может запросить повторную оценку у любой другой рейтинговой системы. То же самое происходит, если игру уличат в несоответствии ее контента автоматическому рейтингу – тогда к процессу подключится ESRB, PEGI или другая организация.

Steam не имеет собственной рейтинговой системы и не участвует в других (национальных или

международных) системах. Правила пользования магазином содержат только одно ограничение – им не могут пользоваться лица моложе 13 лет. Поэтому разработчики должны сами указать рейтинг игры на основе правил, которые применяются в странах продажи. Также Steam позволяет установить при размещении игры предупреждающие теги о потенциально неприемлемом контенте.

Свои требования к контенту есть у Apple App Store, а контролирует соблюдение правил поддержка компании. При публикации игры разработчикам нужно выбрать одну из указанных категорий, учитывая особенности контента в проекте. При этом Apple сопоставляет свои возрастные категории с рейтингами ESRB и PEGI.

На платформе могут встречаться ограничения, которые также влияют на возрастной рейтинг, например, внешние ссылки или использование исторической символики. Но нужно учитывать, что внутренние правила магазинов и сервисов не отменяют требований законов страны, граждане которой получают доступ к игре. Если в национальном законодательстве предусмотрены более жесткие критерии для выбора возрастного рейтинга, то разработчик может выбрать отдельный территориальный рейтинг. Важно понимать, что требования рейтинговых систем и законодательства относятся не только к «внутреннему» контенту игры, но и к ее рекламной кампании: площадкам для размещения, содержанию рекламы, часам показа роликов и их аудитории.

Практически все системы возрастных рейтингов разработаны в соответствии с международными нормами, защищающими права детей на безопасность от информации, наносящей им психический, физический и нравственный ущерб.

II. ТРЕБОВАНИЯ РЕЙТИНГОВЫХ СИСТЕМ ПО ВОЗРАСТНЫМ ГРУППАМ

Следует отметить, что разделение аудитории по возрастам может отличаться от системы к системе, однако практически все из них используют стандартные возрастные диапазоны, в частности 3+, 6+, 12+, 16+ и 18+. Стоит обратить внимание, что некоторые рейтинговые системы имеют специфические возрастные ограничения, которые необходимо учитывать если вы планируете выйти на рынки стран, для которых она предназначена. Например, в RARS существует рейтинг 0+, хотя большинство систем устанавливают рейтинг минимум с 3 лет, а в ESRB установлено разделение на рейтинги 10+ и 13+, хотя большинство систем не делит аудиторию с 6 до 12 лет по возрастам. Дополнительно следует отметить специфичность рейтинга «Everyone» в ESRB, на текущий момент этот рейтинг выставляется для целевой аудитории с 3 лет, однако до 2018 года он распространялся только на возраст старше 6 лет.

При подборе целевой аудитории для гейм-дизайнеров существуют два основных параметра, по которым происходит разделение игроков, – это пол и возраст. Однако так как рейтинговые системы пол не учитывают, то цензурные требования будут формироваться только для стандартных (для геймдизайна) возрастных групп [5]. Ниже перечислены требования и ограничения для игр с разделением по возрастам для целевой аудитории.

0-6 лет (PEGI 3; ESRB EC («Early childhood»), E («Everyone»); RARS 0+) – Подходит для любых возрастных групп. Игра может содержать очень условное насилие в комическом подтексте без нанесения вреда мультипликационным персонажам. Игра может содержать оправданные её жанром и (или) сюжетом эпизодические ненатуралистические изображения или описание физического и (или) психического насилия (за исключением сексуального насилия) при условии торжества добра над злом и выражения сострадания к жертве насилия и (или) осуждения насилия. Ребёнок не должен ассоциировать героев с реальными личностями, они должны быть полностью придуманными. Не должно быть никаких звуков или изображений, которые могут испугать маленьких детей. Нет грубой лексики. Не содержит материалов, которые родители могли бы счесть неподходящими. Продукты, получившие данный рейтинг, изначально разрабатываются для детей и обычно представляют собой развивающие игры.

7-9 лет (PEGI 7; ESRB E («Everyone»); RARS 6+) – Насилие, демонстрируемое в проектах с этим рейтингом условно и нереалистично, в основном «мультяшного» характера, с содержанием разрушений и взрывов. Могут присутствовать сцены, способные напугать очень маленьких детей, очень лёгкие выражения. Из вредных веществ допускаются алкоголь и табак, но без прямого употребления. Сексуальное содержание не допускается. Допускаются кратковременные и ненатуралистические изображения или описание заболеваний человека (за исключением тяжелых заболеваний) и (или) их последствий в форме, не унижающей человеческого достоинства. Возможно использование ненатуралистических изображений или описаний несчастного случая, аварии, катастрофы либо ненасильственной смерти без демонстрации их последствий, которые могут вызывать у детей страх, ужас или панику. Игры с таким рейтингом не должны побуждать к совершению антиобщественных действий и (или) преступлений; можно показывать эпизодические изображения или описания этих действий и (или) преступлений при условии, что не оправдывается их допустимость и выражается отрицательное, осуждающее отношение к лицам, их совершающим.

10-13 лет (PEGI 7; ESRB E10+ («Everyone 10 and older»); RARS 6+) – за счет перекрытия рейтингов по возрастным группам требования обычно такие же как для группы 7-9 лет, но при заявке только на рейтинг ESRB возможно расширение контента – Проекты с данным рейтингом могут содержать больше мультипликационного или мягкого насилия, или несколько откровенные сцены, или минимальное количество крови.

13-18 лет – из-за требований рейтинговых систем этот возрастной диапазон имеет смысл делить на 2 интервала (13-16 и 16-18).

13-16 лет (PEGI 12; ESRB T («Teen»); RARS 12+) – Допускается так или иначе выраженное графически насилие по отношению к фантастическим героям (как правило, не схожим с человеком) или неграфическое (без визуализации крови и ранений) – к узнаваемым животным. Дополнительно допускается некоторая графическая нагота. Любая ненормативная лексика должна быть смягчена и не относиться к сексуальным ругательствам. Проекты из данной категории могут содержать насилие, непристойные сцены, грубый юмор, в меру откровенное сексуальное содержание, кровь или нечастое использование ненормативной лексики. Сцены насилия возможны только без натуралистического показа процесса лишения жизни или нанесения увечий при условии, что выражается сострадание к жертве и (или) отрицательное, осуждающее отношение к жестокости, насилию. Игры не должны побуждать к потреблению алкогольной и спиртосодержащей продукции, участию в азартных играх, занятию бродяжничеством или попрошайничеством. Такой возрастной рейтинг RARS могут получить только игры, не эксплуатирующие интереса к сексу и не носящие возбуждающего или оскорбительного характера эпизодические ненатуралистические изображение или описание половых отношений между мужчиной и женщиной, за исключением изображения или описания действий сексуального характера (это требование распространяется только на РФ и некоторые региональные рейтинги).

16-18 лет (PEGI 16; ESRB M («Mature»); RARS 16+) – Игры данного класса могут включать умеренные сцены насильственных действий и сцены сексуального характера, которые должны соответствовать тем, которые могут встретиться в реальной жизни. Допускаются сцены курения табака или принятия наркотиков, отсылки к азартным играм, грубый юмор и бранные слова. Проекты с рейтингом ESRB M («Mature») (17+) могут содержать жестокое насилие, большое количество крови с расчленением, непристойные сексуальные сцены или грубую ненормативную лексику. Допускается изображение или описание

несчастливого случая, аварии, катастрофы, заблуждения, смерти без натуралистического показа их последствий, которые могут вызывать у подростков страх, ужас или панику. Самые жесткие требования накладывает RARS: полностью запрещена нецензурная брань, за исключением отдельных слов и выражений; изображение или описание насилия, преступлений, употребления запрещенных веществ и т.п. возможно только в случае наличие в игре механизмов или способов, отражающих порицание и недопустимость таких действий.

Старше 18 лет (PEGI 18; ESRB AO («Adults Only 18+»); RARS 18+) – Содержание игры только для взрослых старше 18 лет. Продукты из данной категории могут содержать длительные сцены жестокого насилия и/или очень откровенное сексуальное содержание, а также все то, что попадает под запрет для всех других возрастных категорий.

III. ЗАКЛЮЧЕНИЕ

Наличие рейтинга и соблюдение требований рейтинговых систем принципиально важно, если вы планируете выходить на крупные рынки или работать с известными издателями. Однако нужно учитывать, что это не гарантия возможного распространения игры по всему миру. До сих пор существует ряд стран и регионов, которые имеют очень жесткие рейтинговые системы и запреты. Например, большая часть игр с рейтингом 16+ и 18+ в Китае будет подвергаться дополнительной цензуре и разработчикам придется убирать некоторые изображения, элементы и сцены (так произошло в частности с Battlefield 4, Fortnite, Tom Clancy's Rainbow Six: Siege и даже некоторыми настольными карточными играми). Поэтому при разработке игры под конкретную целевую аудиторию необходимо обязательно учитывать требования мировых и региональных рейтинговых систем.

СПИСОК ЛИТЕРАТУРЫ

1. Pan European Game Information [Electronic resource] / PEGI, 1996. – Mode of access: <http://www.pegi.info>. – Date of access: 14.10.2023.
2. Entertainment Software Rating Board [Electronic resource] / Rating Guide - Entertainment Software Association, 1996. – Mode of access: <http://https://www.esrb.org/ratings-guide/>. – Date of access: 14.10.2023.
3. Федеральный закон РФ №436-ФЗ "О защите детей от информации, причиняющей вред их здоровью и развитию" в действующей редакции от 31.07.2020 №03-ФЗ.
4. Шелл, Дж. Геймдизайн. Как создать игру, в которую буду играть все / Дж. Шелл // Издательство: Альпина Паблишер, 2022. – 640 с.

МОДЕЛИРОВАНИЕ ТЕПЛОВЫХ ГРАДИЕНТОВ В СВЧ ПЛАТЕ МАКЕТА АВТОМОБИЛЬНОГО РАДАРА

Журавлёв В. И., Наумович Н. М., Стежко И. К.

Центр 1.6, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: vadh@bsuir.by

Рассмотрено построение тепловой модели для оценки распределения температуры в печатных модулях автомобильных радаров. Результаты моделирования показывают возникновение значительных градиентов температуры, влияющих на выходные параметры.

ВВЕДЕНИЕ

Использование компактных фазированных решёток миллиметровых волн быстро развивается в таких областях как автомобильные радары, спутниковая и транзитная связь, сканирование систем безопасности и формирование изображений. Технология «антенна в корпусе» (AiP) стала основным вариантом интегрированного корпуса с антенной и приёмопередатчиком для применения в автомобильных радаров на частоте от 72 ГГц [1]. При её использовании может быть достигнуто значительное уменьшение габаритных размеров, и уменьшение потерь и повышение эффективности передатчика. Однако для применения в диапазоне миллиметровых волн требует применения дополнительных мер по отводу тепла. Это обусловлено возникновением градиентов температуры в ограниченном пространстве вследствие использования многослойной компоновки печатных плат с теплонагруженными компонентами и технологии AiP. Анализ тепловых характеристик позволяет оптимизировать конструкцию корпуса и систему охлаждения.

1. ТЕПЛОВАЯ МОДЕЛЬ МНОГОСЛОЙНОЙ СТРУКТУРЫ

Поскольку технология AiP на одной плате с приёмопередатчиком представляет собой неоднородную структуру, включающую компоненты с разными тепловыми свойствами. Задача представляет собой процесс распространения теплоты в анизотропном теле, содержащим некоторое количество слоёв. Имеется некоторая сложность получения аналитического решения для многослойной структуры и громоздкости применяемого математического аппарата, поэтому требуется применить некоторые допущения.

В используемых на практике моделях полагают, что слои, сформированные в объёме кристалла, имеют одинаковые теплофизические параметры и различаются только по электрическим характеристикам. Ряд моделей позволяет определить закон теплопереноса в структурах со строго ограниченным количеством слоёв, другие модели рассматривают только процесс постоянного нагрева [2]. В рассматриваемом случае достаточно

определить закон распространения избыточной температуры T в заданном слое структуры при воздействии импульсного теплового источника.

Рассмотрим модель многослойной структуры, на поверхности которой действует импульсный источник тепла P_0 (рис. 1). Каждый слой однороден, но имеет собственное постоянное значение теплопроводности K .

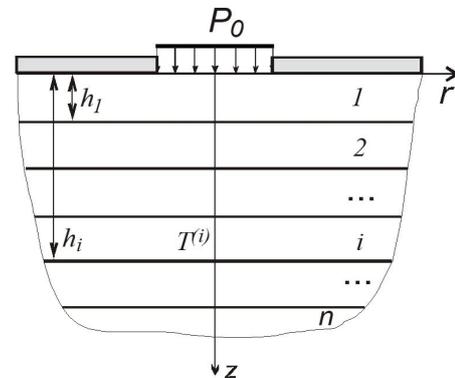


Рис. 1 – Модель многослойной структуры с импульсным источником тепла на поверхности

Двумерное уравнение теплопереноса в структуре можно записать в виде:

$$K_{11} \frac{\partial^2 T}{\partial r^2} + 2K_{12} \frac{\partial^2 T}{\partial r \partial z} + K_{22} \frac{\partial^2 T}{\partial z^2} = 0 \quad (1)$$

Выражение (1) представляет собой линейное дифференциальное уравнение с постоянными коэффициентами. Для i -слоя структуры можно выполнить линейное преобразование координат [3]:

$$\begin{bmatrix} R \\ Z \end{bmatrix} = \begin{bmatrix} 1 & a_i \\ 0 & b_i \end{bmatrix} \begin{bmatrix} r \\ z \end{bmatrix} + \sum_{k=1}^{i-1} h_k \begin{bmatrix} a_i - a_{i+1} \\ b_i - b_{i+1} \end{bmatrix}, \quad (2)$$

где

$$a_i = -\frac{K_{12}^{(i)}}{K_{22}^{(i)}}, \quad b_i = \frac{K_i}{K_{22}^{(i)}}, \quad K_i = \sqrt{K_{11}^{(i)} K_{22}^{(i)} - K_{12}^{(i)^2}}$$

Физический смысл (2) заключается в том, что задача анизотропного теплопереноса представляется эквивалентной задачей изотропного

теплопереноса. Преобразование через (2) является линейным и непрерывным, а также не имеет деформации граничных условий на рассматриваемом участке структуры.

Соответственно,

$$K_{i+1} \left(\frac{\partial^2 T^{(i+1)}}{\partial R^2} + \frac{\partial^2 T^{(i+1)}}{\partial Z^2} \right) + K_1 \left(\frac{\partial^2 T^{(1)}}{\partial R^2} + \frac{\partial^2 T^{(1)}}{\partial Z^2} \right) = 0 \quad (3)$$

Уравнение (3) при заданных граничных условиях можно аналитически решать поэтапно двунаправленными преобразованиями Фурье [4]. Использование численных методов позволяет значительно сократить затраты на поиск решения для произвольного слоя.

II. МОДЕЛИРОВАНИЕ ГРАДИЕНТА В ТЕПЛОТВОДЯЩЕМ СЛОЕ

Моделирование распределение температур в AiP плате макета автомобильного радара (рисунок 2) выполнялось с применением численного метода конечных элементов с открытыми граничными условиями второго рода для нормальных условий эксплуатации при температуре окружающей среды (воздух) 35 °C [5]. Однотипные материалы заменялись аналогичными, но с отличающимися нетепловыми свойствами во избежание пропуска границ тепловых сопротивлений при однородной сетке.

Результаты моделирования для платы приёмопередающего модуля (СВЧ) наиболее важны, так как возникновение больших перепадов температуры в ней наиболее критично. Вид её полученного теплового профиля в составе корпуса приведён на рисунке 3 (корпус не показан).

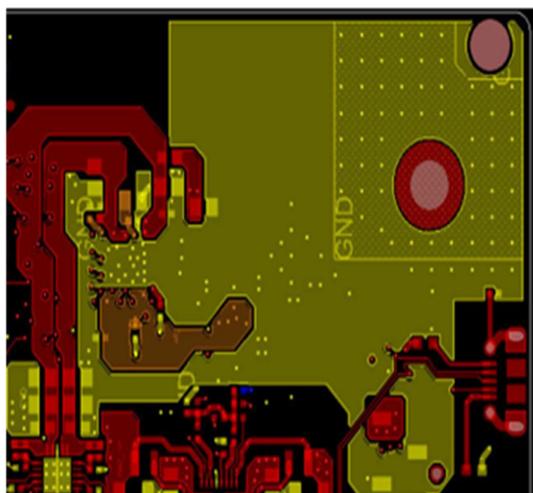


Рис. 2 – Теплоотвод металлизацией на нижней стороне платы СВЧ

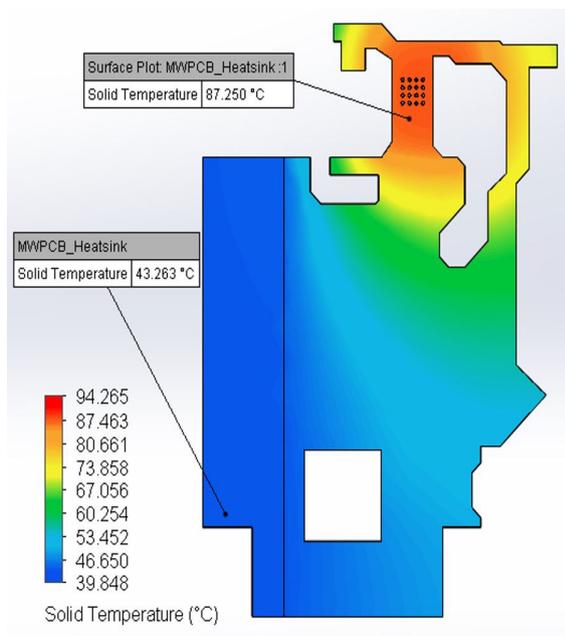


Рис. 3 – Расчётное распределение температуры по теплоотводящему слою

III. ЗАКЛЮЧЕНИЕ

Моделирование теплового профиля в многослойных структурах автомобильных радаров указывает на возникновение довольно больших градиентов температур. Это следует учитывать в дальнейшем при возможном проектировании дополнительного теплоотвода, так как может привести возникновению к термомеханическим напряжениям и ухудшить параметры передающей антенны.

СПИСОК ЛИТЕРАТУРЫ

1. Han, Y. AiP component and board level heat dissipation analysis for automotive radar / Y. Han, T.C. Chai and S.S. H. Lim // 2021 IEEE 71st Electronic Components and Technology Conference (ECTC). – 2021. – P.1850–1857.
2. D'Alessandro, V. A critical review of thermal models for electro-thermal simulation / V. d'Alessandro, N. Rinaldi // Solid-State Electronics. – 2002. – №46. – P.487–496.
3. Ma, C.C. Analytical exact solutions of heat conduction problems for anisotropic multi-layered media / C.C. Ma, S.W. Chang // International J. of Heat and Mass Transfer. – 2004. – №47. – P.1643–1655.
4. Dwyer, V.M. Thermal failure in semiconductor devices. / V.M. Dwyer, A.J. Franklin, D.S. Campbell // Solid-State Electronics. – 1990. – Vol.33, №5. – P. 553–560.
5. Журавлёв, В.И. Моделирование теплового режима образца автомобильного радара / В.И. Журавлёв, Н.М. Наумович, С.А. Корневский // «Приборостроение-2022»: материалы 15-й Международной научно-технической конференции, – Минск : БНТУ, 2022. – С.253-254.

МОДЕЛИРОВАНИЕ ВЗАИМНОГО РАЗОГРЕВА КОМПОНЕНТОВ СИЛОВЫХ ПРЕОБРАЗОВАТЕЛЕЙ

Журавлёв В. И., Петровский И. И., Свито И. Л.

Кафедра теоретических основ электротехники, Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: vadzh@bsuir.by

Предложен подход для оценки взаимного разогрева полупроводниковых компонентов силовых преобразователей на основе теплового моделирования. Показано отличие в получаемых результатах в зависимости от учёта параметров компонентов преобразователя

ВВЕДЕНИЕ

В последние несколько лет возрос интерес к оценке надёжности силовых электронных устройств на основе физики отказов [1]. Электронные полупроводниковые ключи являются наиболее уязвимыми частями силовых преобразователей и, следовательно, играют важную роль в их надёжности. Механизм термомеханического разрушения имеет большое значение из-за скачков температуры перехода и разности коэффициентов теплового расширения различных слоёв [2]. При выделении тепловой мощности на различных компонентах на плате между ними могут возникнуть термомеханические напряжения. Это проявляется в увеличении напряжения включения и повышении теплового сопротивления компонентов [3]. Как результат, это сказывается на выходных параметрах всего преобразователя и приводит к ухудшению электрических и тепловых характеристик эксплуатации, что приводит к ускоренной деградации. Эти процессы могут быть спрогнозированы посредством построения соответствующих моделей теплового моделирования.

I. ТЕПЛОВАЯ МОДЕЛЬ

При построении тепловой модели важно учитывать взаимное влияние рядом находящихся компонентов в конструкции преобразователя. Поэтому строится САД-модель преобразователя с использованием геометрической информации о структуре компонентов. Далее задаются термофизические параметры структуры и производится решение методом конечных элементов. В качестве исходной структуры рассматривается схема преобразователя, содержащего ключевой IGBT транзистор и диод, имеющих общую систему охлаждения (рис. 1).

Наиболее информативным параметром оценки нагрева электронного компонента является температура его самой нагретой части – температура перехода. Оценить температуру перехода (самой нагретой зоны кристалла) каждого из компонентов можно из функциональных уравнений, связывающих температуру окружающей среды, тепловое сопротивление «переход

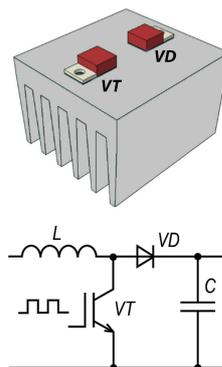


Рис. 1 – Компоненты преобразователя

– окружающая среда», тепловое сопротивление «переход – печатная плата» (основание), среднюю температуру платы непосредственно около корпуса ключа с его рассеиваемой мощностью. В дальнейшем оценивается тепловое взаимодействие транзистора и диода.

Для получения решения необходимо рассчитать рассеиваемую мощность, которая используется в качестве источника тепла в тепловой модели. Потери мощности могут быть рассчитаны с помощью термоэлектрической модели в SPICE, учитывающей электрическое поведение, а значение избыточной рассеиваемой мощности может быть получено из характеристик тепловых потерь мощности компонента [4]. Для извлечения параметров модели и определения теплового сопротивления между двумя рассматриваемыми компонентами на основе температурных откликов в зависимости от времени, полученных с помощью метода конечных элементов [5].

II. МОДЕЛИРОВАНИЕ ВЗАИМНОГО РАЗОГРЕВА

Для получения значений теплового импеданса в рамках предлагаемого метода были проведены два переходных моделирования, в которых в качестве единственного источника тепла использовались потери мощности либо IGBT, либо диода. Для упрощения САД-модели компоненты установлены на ребристом радиаторе с направленным воздушным охлаждением с возможностью комбинации попеременного задания теплового источника (рис.3). При моделировании на рёбра радиатора накладывается граничное условие

конвективного теплообмена, коэффициент теплоотдачи составляет 35 (Вт*м/К), а температура окружающей среды задаётся равной 20°C. Выделяемая мощность в IGBT транзисторе и диоде составила примерно 11 Вт и 7 Вт соответственно.

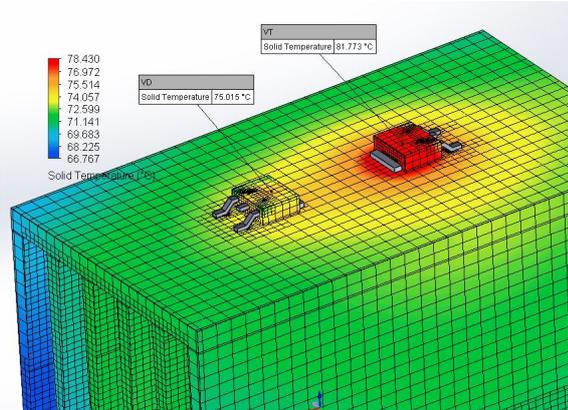


Рис. 2 – Упрощённая тепловая модель разогрева

Для анализа влияния нелинейности и взаимной тепловой связи между компонентами проанализированы четыре варианта модели: линейная модель без связи, в которой электрические свойства не зависят от температуры; нелинейная модель без связи; нелинейная связанная, в которой учитывается тепловая связь компонентов и нелинейная не связанная, в которой связь не учитывается. При анализе связанной модели учитывается взаимный импеданс нагрева, в то время как в несвязанной модели взаимный импеданс устанавливается равным нулю. Сравнение температурных профилей показано на рис. 3. Графики 1 и 2 показывают температуру перехода IGBT транзистора и диода без учёта взаимной тепловой связи соответственно. Эти значения существенно отличаются от результатов для вариантов модели, когда взаимный разогрев учитывается (графики 3 и 4).

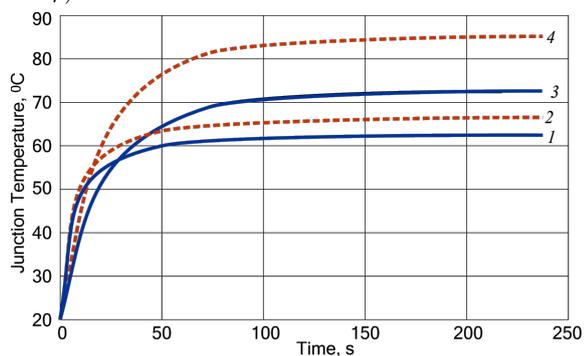


Рис. 3 – Рассчитанные тепловые профили

Прогнозируемые температуры варианты моделей с тепловой связью и моделей без связи существенно различаются. Модель без учёта взаимного нагрева показывает меньшую температуру перехода диода и IGBT транзистора на 12°C и 25°C соответственно по сравнению с моделью со связью. Полная температура перехода транзистора или диода складывается из вклада от саморазогрева, взаимного разогрева и температуры окружающей среды. Помимо наличия значительного теплового градиента, такая разница может быть причиной возникновения термомеханических напряжений в корпусах приборов и при их размещении на печатной плате.

Однако в зависимости от применяемого способа охлаждения взаимное влияние друг на друга полупроводниковых приборов оказывается весьма значительным. В результате моделирования потерь и теплового режима можно сделать вывод о том, что на срок службы действительно влияет взаимное влияние синхронности работы управляемых и неуправляемых ключей преобразователя.

Предложенный подход позволяет получить более релевантные результаты, чем при расчётах, основанных на единичных параметрах, полученных из спецификации производителя.

СПИСОК ЛИТЕРАТУРЫ

1. Wang, H., et al. Transitioning to Physics-of-Failure as a Reliability Driver in Power Electronics // IEEE Journal of Emerging and Selected Topics in Power Electronics. – 2014. – vol. 2, no. 1. – P. 97-114.
2. Chen, J. et al. E. Impact of load pulse duration on failure mechanism of high power IGBT modules under power cycling condition / The 16th IET International Conference on AC and DC Power Transmission (ACDC 2020). – 2020. – P. 412-417.
3. Durand, C., et al. Power Cycling Reliability of Power Module: A Survey // IEEE Transactions on Device and Materials Reliability. – 2016. – vol. 16, no. 1. 2016. – P. 80-97.
4. Shahjalal, M. Electro-thermal modelling of multichip power modules for high power converter application / M. Shahjalal, H. Lu, C. Bailey / 2017 18th International Conference on Electronic Packaging Technology (ICEPT), Harbin, China, 2017. – P. 940-945.
5. Bahman, A.S. et al. A 3-D-Lumped Thermal Network Model for Long-Term Load Profiles Analysis in High-Power IGBT Modules // IEEE Journal of Emerging and Selected Topics in Power Electronics. – 2016.– vol. 4, no. 3. – P. 1050-106.

МОДЕЛИРОВАНИЕ РЕЗИНО-МЕТАЛЛИЧЕСКИХ ВИБРОИЗОЛЯТОРОВ ДЛЯ СИНТЕЗАТОРОВ ЧАСТОТ

Кийко В.Н., Ахапкина А. М., Колбун В. С.

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {kiyko}@bsuir.by

В данной статье рассматривается виброзащита кварцевых генераторов в конструкциях синтезаторов частот с применением резино-металлических виброизоляторов. Проведён гармонический анализ при воздействии внешних вибраций. В результате моделирования определены резонансные частоты для учёта в проектировании конструкций синтезаторов частот при защите от случайной вибрации и ударных нагрузок.

ВВЕДЕНИЕ

Одним из ключевых компонентов современных радиоэлектронных систем являются высокостабильные виброустойчивые синтезаторы частот [1]. Эти системы размещаются на подвижных объектах: самолетах, вертолетах, кораблях, беспилотных летальных аппаратах, автомобильной и гусеничной технике, ракетах, спутниках – и при работе подвергаются воздействию механической вибрацией, которая влияет на частоту и фазу генерируемого сигнала [2]. Приоритетной задачей является защита от вибрации кварцевых генераторов, как оказывающих наибольшее влияние на фазовые характеристики синтезаторов частот [3].

I. МОДЕЛИРОВАНИЕ РЕЗИНОМЕТАЛЛИЧЕСКОГО ВИБРОИЗОЛЯТОРА

При создании параметрической модели виброизолятора использовалась среда для трёхмерного автоматизированного проектирования деталей и сборок. Такая модель состоит из трёхмерной геометрии твердотельных элементов. Сначала рисуется эскиз, создается основание, а затем в модель добавляются многочисленные элементы. На рисунке 1 представлена модель такого виброизолятора.

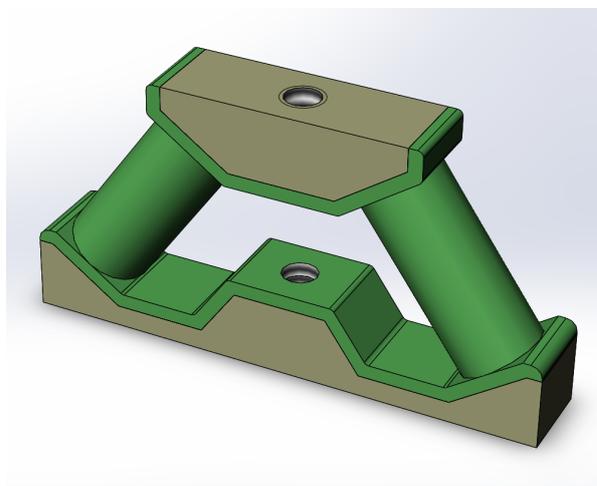


Рис. 1 – Модель резино-металлического виброизолятора

Численно виброизолятор характеризуется масс-габаритными характеристиками, коэффициентом жёсткости, зависимостью резонансной частоты от поддрессоренной массы.

Их особенностью являются способность выдерживать довольно большое удлинение, что эффективно для демпфирования не только для вибрации, но и ударов. Для их изготовления используется высококачественный каучук или силикон с высоким затуханием.

Моделирование вибрационных характеристик резино-металлического виброизолятора проводили с помощью численного метода конечных элементов [4], в основу которого заложены два типа исследований: статический и модальный. При статическом типе выполняются анализ жесткостей моделей виброизоляторов в различных направлениях, зависящих от упругих или вязкоупругих свойств материалов, а также сопоставление полученных результатов расчета со значениями, определенными производителем в соответствующих каталогах. Модальный тип исследований предназначен для определения собственных форм и частот колебаний, которые, в свою очередь, являются резонансными для виброизолятора.

Начальные и граничные условия:

- основание модели жестко закреплено;
- на подвижную часть модели приложена сосредоточенная сила равная 10 Н.

На рисунке 2 представлена расчётная модель резино-металлического виброизолятора. Результаты статического анализа представлены на рисунке 3 в виде картины распределения перемещений.

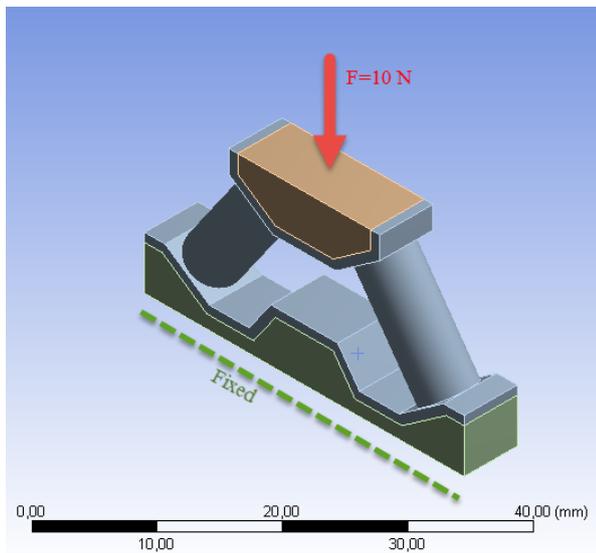


Рис. 2 – Расчётная модель резино-металлического виброизолятора

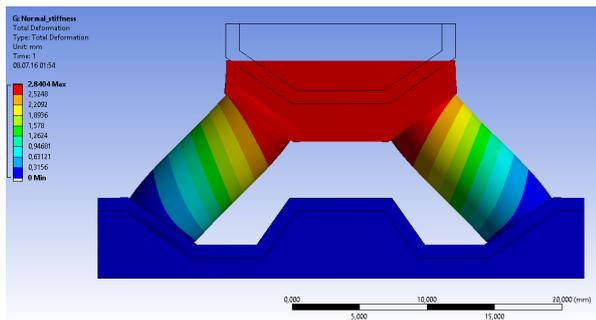


Рис. 3 – Распределение виброперемещений резино-металлического виброизолятора

Максимальное перемещение составляет 2,84 мм, что соответствует осевой жёсткости 3,52 Н/мм.

Граничными условиями для модального анализа является только жёсткое закрепление основания модели. На рисунке 4 приведены первые шесть форм собственных частот и их значения.

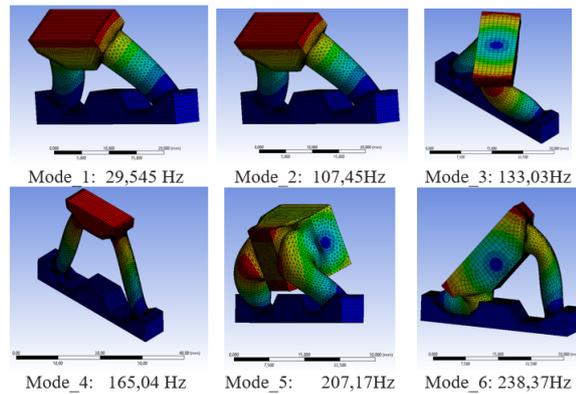


Рис. 4 – Формы собственных частот резино-металлического виброизолятора

Модальный анализ позволяет оценить возможность появления резонансных режимов, которые могут возникнуть при попадании собственных частот в рабочий диапазон действующих внешних вибраций. На основании этих данных можно предусмотреть дополнительную защиту на этих частотах и оптимизировать конструкцию для повышения виброустойчивости синтезатора частот.

II. ЗАКЛЮЧЕНИЕ

В качестве дальнейшего развития приведённого метода исследования необходимо моделирование большего количества средств уменьшения вибрации и материалов для оптимизации конструкции (коэффициенты демпфирования и вязкоупругие характеристики материалов требуется уточнять экспериментальным способом). Предварительный расчёт резино-металлического виброизолятора и моделирование его использования в составе синтезатора частот позволяют определить оптимальные характеристики его конструкции, исходя из спектра воздействующей вибрации и параметров кварцевого генератора.

СПИСОК ЛИТЕРАТУРЫ

1. Муравьев В.В. Синтезатор СВЧ с прямым цифровым синтезом // Приборы и методы измерений. 2014. №1 (8). С. 52-56.
2. Карпушин В. Б. Вибрации и удары в радиоаппаратуре. М.: Сов. радио, 1971. – 344 с.
3. Apte A., Rohde U. L., Poddar A., Rudolph M. Optimizing Phase-Noise Performance: Theory and Design Techniques for a Crystal Oscillator // IEEE Microwave Magazine. 2017. V. 18. N. 4. P. 108-123.
4. Pawel Kurowski, "Chapter 11: Vibration Analysis," in Finite Element Analysis for Design Engineers, SAE, 2017, pp. 181-196.

ИНТЕГРАЛ ВЕРОЯТНОСТЕЙ ОШИБКИ ПРИ ИДЕНТИФИКАЦИИ ВЫБОРКИ ЧИСЛОМ ВП

Кобяк И. П.

Кафедра электронных вычислительных машин,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: ipkobayak2012@mail.ru

В представляемой работе на основе конечных разностей получено численное значение, характеризующее нижнюю границу энумератора для вероятности пропуска ошибки при наблюдении векторов переходов в последовательностях со случайной природой. Рассмотрен частный случай производящей функции для одной переменной, характеризующей близкое к точному значение параметра в асимптотике. Получено соотношение, характеризующее точку моды в функции плотности распределения вероятностей ошибки при регистрации сложных событий или объектов на основе векторов переходов.

ВВЕДЕНИЕ

Многомерные векторы переходов (ВП) заданного вида формируются поразрядным дифференцированием статических векторов случайного процесса с последующим поразрядным однопериодным сглаживанием отрицательных событий. Для регистрации схемами наблюдения каналов связи выбираются однотипные ВП, образующие лебеговскую меру событий во множестве всех векторов переходов. Основным достоинством принципа формирования точечных оценок из таких субдинамических объектов является низкая вероятность пропуска ошибки, что следует из факта сложности самих объектов и невозможности (при объединении вероятностей) получения полной группы событий. Особый интерес при сравнении данного метода с классическими алгоритмами свертки вызывает механизм вычисления площади под кривой графика плотности распределения вероятностей при наблюдении ВП. Поставленная задача в данной работе решается путем суммирования факториальных моментов, входящих в состав перечисляющей производящей функции, на основе алгоритма конечных разностей.

I. ВЕРОЯТНОСТЬ ПРОПУСКА ОШИБКИ И ПРОИЗВОДЯЩАЯ ФУНКЦИЯ

Из результатов работы [1] следует, что гладкая часть функции нормированной мощности классов эквивалентности (МКЭ), соответствующей вероятности пропуска ошибки при регистрации ВП в частном случае при $j = 1$ может быть представлена соотношением вида:

$$P_{ifc}(gl) = \sum_g \pi(g) \sum_{n-g} \frac{1}{m^n} (3^{r-\mu})^{\sum_{i=1}^{n-4} k_{1,i} + k_{1,n-2}} \times \\ \times m^{\sum_{i=1}^{n-4} i k_{1,i} + (n-2)k_{1,n-2}} \times \prod_{i=2}^{n-4} \left(\frac{1}{2^{i+1}} \beta_{1,i} \right)^{k_{1,i}} \times \\ \times \left(\frac{1}{2^{n-1}} \beta_{1,n-2} \right)^{k_{1,n-2}} \frac{(n-g)!}{k_{1,1}! k_{1,2}! \dots k_{1,n-4}! k_{1,n-2}!},$$

где ортонормирующие значения $g = 1 + i$ берутся из разбиений параметра

$$n = \sum_{i=1}^{n-4} (2+i)k_{1,i} + nk_{1,n-2}$$

на целые части вида $2+i$, при $k_{1,n-3} = 0$; $\pi(g) = \pi(n) -$ это композиция текущего разбиения числа n . Причем $n-g = \sum_{i=1}^{n-4} k_{1,i} + k_{1,n-2}$, а параметр

$$\beta_{1,i} = \frac{1}{\xi} \left[(1+\xi)^{i+1} - (1-\xi)^{i+1} \right] \\ \xi = \sqrt{1-4p}.$$

Значение вероятности p в данном равенстве определено как

$$p = \frac{3^{r-\mu}}{m^2}.$$

Производящая функция для вероятности ошибки $P_{ifc}(gl)$ может быть получена путем преобразования соответствующего произведения факториальных моментов в сумму с использованием полиномиального закона. При этом имеем:

$$P_{ifc}(gl) = \sum_g \pi(g) \left(k_{1,1} p e^t x_1^1 + \sum_{i=2}^{n-4} k_{1,i} \frac{p e^t x_1^i}{2^{i+1}} \beta_{1,i} + \right. \\ \left. + k_{1,n-2} \frac{p e^t x_1^{n-2}}{2^{n-1}} \beta_{1,n-2} \right)^{n-g}.$$

С учетом моды данное равенство представимо в виде

$$P_{ifc} = \frac{1}{2^{n+1}} \beta_n + P_{ifc}(gl).$$

Используя алгоритм конечных разностей определим частное значение площади под кривой распределения функции вероятности ошибки при наблюдении лебеговской меры ВП.

II. НИЖНЯЯ ГРАНИЦА ДЛЯ ИНТЕГРАЛА ВЕРОЯТНОСТЕЙ ОШИБКИ

Достаточно просто показать, что максимальное значение вероятности p для всех значений $r > 1$ равно $\frac{3}{16}$, что следует из тривиальных подстановок параметров r и μ . Использование данного значения в последующих вычислениях

обусловлено тем фактом, что для различных значений p конечные соотношения в процессе вывода представляются множеством разветвленных и трудно формализуемых равенств. В связи с этим рассмотрим разность двух аппроксимированных факториальных моментов $\beta_{1,j}$ и $\beta_{1,j+1}$, определяющих функцию правдоподобия в вероятности ошибки, при $j = 1$ и произвольных значениях i :

$$\begin{aligned} & \left(\frac{1}{2^{i+1}} \beta_{1,i} - \frac{1}{2^{i+2}} \beta_{1,i+1} \right) = \\ & = \frac{1}{2^{i+1}} \frac{1}{\xi} \left[(1 + \xi)^{i+1} - (1 - \xi)^{i+1} \right] - \\ & \quad - \frac{1}{2^{i+2}} \frac{1}{\xi} \left[(1 + \xi)^{i+2} - (1 - \xi)^{i+2} \right]. \end{aligned}$$

Данная разность достаточно просто преобразуется к виду:

$$\begin{aligned} & \frac{1}{\xi} \left[A^{i+1} (1 - A) - B^{i+1} (1 - B) \right] = \\ & = 2 \left(A^{i+1} \frac{1}{4} - B^{i+1} \frac{3}{4} \right) = \frac{3}{2} \left(\frac{1}{4} \right)^{i+1} (3^i - 1), \end{aligned}$$

где

$$A = \left(\frac{1}{2} + \frac{1}{2} \xi \right), \quad B = \left(\frac{1}{2} - \frac{1}{2} \xi \right).$$

Для формирования интегральной суммы будем использовать концепцию суммирования вероятностей с одномоментным их представлением в интеграле, то есть при $\pi(g) = 1$. При этом необходимо решать задачу суммирования рядов, что предполагает введение нового параметра для длины фрагмента выборки $n_1 \gg 1$, но $n_1 \ll n$.

Рассмотрим теперь полученную выше производящую функцию в виде

$$P_{ifc} = Mo + \left[p \left(1 + \sum_{i=2}^{n_1-4} \frac{1}{2^{i+1}} \beta_{1,i} + \frac{1}{2^{n_1-1}} \beta_{1,n_1-2} \right) \right]^n.$$

В приведенном соотношении сумма элементов в скобках, выраженная с использованием полученных выше разностей, может быть преобразована в соотношение:

$$\begin{aligned} & \sum_{i=2}^{n_1-4} \frac{1}{2^{i+1}} \beta_{1,i} = (n_1 - 5) \frac{1}{2^{2+1}} \beta_{1,2} - \\ & - \frac{n_1-3}{2} \sum_{i=3}^{n_1-4} \left(\frac{3}{4} \right)^i + \frac{1}{2} \sum_{i=3}^{n_1-4} i \left(\frac{3}{4} \right)^i + \\ & + \frac{3}{2} (n_1 - 3) \sum_{i=3}^{n_1-4} \left(\frac{1}{4} \right)^i - \frac{3}{2} \sum_{i=3}^{n_1-4} i \left(\frac{1}{4} \right)^i. \end{aligned}$$

Для решения поставленной задачи выполним суммирование дробей в данном соотношении при $n_1 - 4 \gg 1$. Тогда для простых рядов имеем:

$$\sum_{i=3}^{n_1-4} \left(\frac{3}{4} \right)^i = \frac{27}{16}, \quad \sum_{i=3}^{n_1-4} \left(\frac{1}{4} \right)^i = \frac{1}{48}.$$

Для сложных рядов будем использовать интегрально-дифференциальные преобразования. При этом:

$$\sum_{i=3}^{n_1-4} i \left(\frac{3}{4} \right)^i = \frac{3}{4} \sum_{i=1}^{\infty} i \left(\frac{3}{4} \right)^{i-1} - \frac{15}{8}.$$

Интегрируя сумму ряда в данном равенстве, имеем:

$$\int \sum_{i=3}^{n_1-4} i \left(\frac{3}{4} \right)^i di = \frac{3}{4} \int \sum_{i=1}^{\infty} i \left(\frac{3}{4} \right)^{i-1} = \frac{3}{4} \left(1 - \frac{3}{4} \right)^{-1}.$$

Продифференцируем полученное значение в виде:

$$\frac{3}{4} \frac{\partial}{\partial t} (1-t)^{-1} = \frac{3}{4} \left(1 - \frac{3}{4} \right)^{-2}.$$

Отсюда окончательно имеем:

$$\sum_{i=3}^{n_1-4} i \left(\frac{3}{4} \right)^i = \frac{3}{4} \left(1 - \frac{3}{4} \right)^{-2} - \frac{15}{8} = \frac{81}{8}.$$

Аналогично можно вычислить сумму и оставшихся членов ряда:

$$\sum_{i=3}^{n_1-4} i \left(\frac{1}{4} \right)^i = \frac{1}{4} \left(1 - \frac{1}{4} \right)^{-2} - \frac{3}{8} = \frac{15}{72}.$$

Тогда используемый фрагмент производящей функции может быть представлен в виде:

$$P_{ifc} = Mo + \left[p \left(1 - \frac{13}{8} + \frac{81}{16} - \frac{45}{144} + \frac{3^{n_1-1}}{4^{n_1-1}} \right) \right]^{n \gg n_1}$$

или

$$P_{ifc} = Mo + \left(p \frac{33}{8} \right)^{n \gg n_1} = \frac{3}{2} \left(\frac{3}{4} \right)^n + \left(\frac{99}{128} \right)^n.$$

В асимптотике дроби в степени n очевидно представляют собой весьма малые значения. Таким образом, метод идентификации последовательностей с помощью точечных оценок вероятности наблюдения ВП следует считать наиболее предпочтительным из всех известных.

III. Выводы

Классические алгоритмы свертки цифровых последовательностей, как правило, основываются на функции суммирования выбранных элементарных событий в пересылаемых сообщениях. При этом площадь под огибающей кривой плотности распределения вероятностей принципиально равна единице, что следует из факта полной группы событий для вероятностей наблюдаемых объектов. Рассматриваемый же в данной работе алгоритм являет собой методологию особого вида, основанную на переключении технических систем из состояния в состояние под действием синхрипульсов. Соответственно задача наблюдения лебеговской меры ВП не определяет возможности построения графика вероятностей ошибки для полной группы событий, что и определяет эффективность метода.

1. И.П.Кобяк. Производящая функция для вероятности пропуска ошибки при наблюдении векторов переходов // В сб. науч. статей IX Международной науч.-практич. конференции «BIG DATA and Advanced Analytics», Минск, 17-18 мая 2023 г. Часть 2. – Минск : Бестпринт, 2023. – С. 16–23.

АЛГОРИТМЫ МАСШТАБИРОВАНИЯ ГРАФИЧЕСКИХ ЭЛЕМЕНТОВ

Коршикова Д. В., Кукин Д. П., Купчина Е. В.
Кафедра вычислительных методов и программирования,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {korshikova, kukin, e.kupchina}@bsuir.by

В данной статье будут рассмотрены алгоритмы, которые используются для масштабирования графических элементов, определены достоинства и недостатки каждого из них и специфика их применения.

ВВЕДЕНИЕ

Под масштабированием изображения следует понимать увеличение или уменьшение размера изображения с сохранением его пропорций. Алгоритмы масштабирования графических элементов определяют, каким образом изменять размеры объектов на экране. При этом, в зависимости от типа графики (растровая, векторная), масштабирование производится разными алгоритмами.

I. ЛИНЕЙНОЕ МАСШТАБИРОВАНИЕ

Линейное масштабирование - это процесс изменения размеров объекта путем умножения его координат на определенный масштабный коэффициент. При линейном масштабировании все расстояния в объекте увеличиваются или уменьшаются пропорционально. Математически линейное масштабирование можно выразить следующим образом: $x' = kx$, $y' = ky$, где x' и y' - новые координаты объекта после масштабирования, x и y - исходные координаты объекта, k - масштабный коэффициент. Значение масштабного коэффициента k может быть больше 1, что приведет к увеличению объекта, или меньше 1, что приведет к уменьшению объекта. В компьютерной графике линейное масштабирование может применяться для изменения размера изображений, масштабирования графических элементов или обработки входных данных для лучшего просмотра и анализа. Преимущества линейного масштабирования:

1. Простота реализации: линейное масштабирование является наиболее простым алгоритмом масштабирования и может быть реализован сравнительно легко.
2. Быстрота выполнения: линейное масштабирование требует минимальных вычислительных ресурсов, поэтому работает быстро.

Недостатки линейного масштабирования:

1. Искажение пропорций: при линейном масштабировании объекта или изображения в одном направлении (например, только по горизонтали или только по вертикали), происходит искажение пропорций. Это может привести к деформации объекта или изображения.

2. Потеря качества: линейное масштабирование может приводить к потере деталей и резкости изображения, особенно при увеличении размера. Изображение выглядит размытым и недостаточно четким.
3. Ограниченные возможности: линейное масштабирование предоставляет только базовые возможности масштабирования и не может обеспечить высокую степень гибкости или точности при изменении размера объекта или изображения.
4. Неэффективная обработка точечных объектов: линейное масштабирование может не подходить для точечных объектов, таких как текст или линии, так как они могут стать нечеткими и нечитаемыми после масштабирования. В целом, линейное масштабирование является простым и быстрым алгоритмом, но оно может быть неприемлемым, если требуется высокое качество и сохранение деталей при масштабировании.

II. АФФИННОЕ МАСШТАБИРОВАНИЕ

Аффинное масштабирование - это процесс изменения размеров объекта путем умножения его координат на определенные масштабные коэффициенты, а также добавления смещения. В отличие от линейного масштабирования, аффинное масштабирование позволяет изменять размеры объекта независимо по разным осям. Математически аффинное масштабирование можно выразить следующим образом: $x' = ax + c$, $y' = by + d$, где x' и y' - новые координаты объекта после масштабирования, x и y - исходные координаты объекта, a и b - масштабный коэффициент, c и d - смещения для осей x и y соответственно. Значения масштабных коэффициентов a и b могут быть разными для осей x и y , что позволяет изменять размер по каждой оси независимо. Смещения c и d позволяют сдвигать объект после масштабирования. Аффинное масштабирование применяется в компьютерной графике для изменения размеров и положения объектов, создания эффектов анимации и прочих операций, связанных с преобразованием объектов.

Преимущества аффинного масштабирования:

1. Сохранение пропорций: аффинное масштабирование позволяет сохранять пропорции объекта при изменении его размера, объекты остаются правильной формы и не деформируются.
2. Гибкость: аффинное масштабирование предлагает широкий спектр масштабирования - оно может производить увеличение или уменьшение объекта с большим контролем над размером.
3. Высокое качество: аффинное масштабирование позволяет сохранять больше деталей и резкости изображения при масштабировании, особенно при увеличении размера.

Недостатки аффинного масштабирования:

1. Сложность реализации: аффинное масштабирование является сложным алгоритмом, требующим сложных вычислений по сравнению с линейным масштабированием.
2. Ограниченные возможности масштабирования: аффинное масштабирование может иметь ограниченные возможности в изменении размера объектов.
3. Потребность в ресурсах: аффинное масштабирование может требовать больше вычислительных ресурсов для его проведения, поэтому может быть более медленным по сравнению с более простыми методами.

В целом, аффинное масштабирование предлагает большую гибкость и сохранение пропорций объектов, но требует более сложной реализации и может быть более ресурсоемким. Этот метод лучше подходит для задач, требующих высокого качества и точности масштабирования.

III. МУЛЬТИСКЕЙЛИНГОВОЕ МАСШТАБИРОВАНИЕ

Мультискейлинговое масштабирование – это процесс изменения размеров объекта с помощью нескольких масштабных коэффициентов, применяемых к различным частям объекта. В отличие от линейного или аффинного масштабирования мультискейлинговое масштабирование позволяет изменять размеры объекта с разной степенью масштабирования для разных его частей. В алгоритмическом мультискейлинговом масштабировании объект разделяется на несколько частей или сегментов, для каждого из которых определяются масштабные коэффициенты. Затем каждый сегмент масштабируется независимо от остальных.

Преимущества мультискейлингового масштабирования:

1. Гибкость и адаптивность: мультискейлинговое масштабирование позволяет использо-

вать различные алгоритмы масштабирования для разных уровней масштабирования. Это позволяет лучшим образом учитывать особенности объекта или изображения на каждом масштабе и снижает возможность потери качества.

2. Сохранение деталей: благодаря использованию разных алгоритмов на разных уровнях масштабирования, мультискейлинговое масштабирование может обеспечивать лучшее сохранение деталей и качества изображения. Это особенно заметно при увеличении размера, где изображение может сохранять резкость и четкость.
3. Более эффективное использование ресурсов: мультискейлинговое масштабирование может распределять вычислительные ресурсы более эффективно.

Недостатки мультискейлингового масштабирования

1. Сложность реализации: мультискейлинговое масштабирование требует более сложной реализации, поскольку включает в себя использование различных алгоритмов на разных уровнях масштабирования и их координацию.
2. Дополнительная вычислительная нагрузка: мультискейлинговое масштабирование может быть более ресурсоемким, поскольку требует обработки изображения на нескольких масштабах и применения разных алгоритмов.

В целом мультискейлинговое масштабирование может быть особенно полезным в компьютерной графике и анимации, где позволяет создавать более реалистичные и детализированные эффекты деформации объектов.

IV. ЗАКЛЮЧЕНИЕ

Это только некоторые из алгоритмов масштабирования графических элементов, и каждый алгоритм имеет свои преимущества и недостатки. Выбор определенного алгоритма масштабирования зависит от требований конкретного приложения и доступных ресурсов. Иногда оптимальное решение может быть комбинацией различных алгоритмов масштабирования, чтобы учесть разные аспекты и обеспечить наилучший результат.

СПИСОК ЛИТЕРАТУРЫ

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс // Техносфера – 2012. – С. 1105.
2. Вольберг, Дж. Искажение цифрового изображения / Дж. Вольберг, Н. Н. // Каталог библиотеки – 1992. – С. 318.

АЛГОРИТМЫ ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ

Коршикова Д. В., Кукин Д. П., Купчина Е. В.
Кафедра вычислительных методов и программирования,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {korshikova, kukin, e.kupchina}@bsuir.by

В данной статье будут рассмотрены алгоритмы, которые используются для восстановления изображений.

ВВЕДЕНИЕ

Алгоритмы восстановления изображений используются для восстановления качества изображений, которые могут быть повреждены, зашумленными или искаженными.

I. ВОССТАНОВЛЕНИЕ ПОСРЕДСТВОМ ФИЛЬТРАЦИИ

Восстановление посредством фильтрации относится к процессу удаления нежелательных или поврежденных элементов из сигнала или данных, чтобы восстановить исходное состояние или качество. Это может быть использовано в различных контекстах, например, в обработке сигналов, изображений или звука.

Фильтрация может производиться различными способами, в зависимости от типа данных и требуемых результатов. Например, в обработке цифровых изображений часто используются фильтры размытия или усиления, чтобы улучшить четкость изображения или снизить шум. В процессе восстановления посредством фильтрации обычно применяются математические алгоритмы, которые анализируют и модифицируют сигнал или данные в соответствии с заданным фильтром. Это может включать выполнение операций, таких как свертка, преобразование Фурье или корреляция.

Цель восстановления посредством фильтрации состоит в том, чтобы устранить несовершенства или дефекты в данных и получить более точное или чистое представление исходного сигнала. В результате может быть улучшена разборчивость, устранена помеха или восстановлена детализация, в зависимости от заданных критериев фильтрации. Восстановление посредством фильтрации может быть полезным во многих областях, включая медицинскую диагностику, обработку изображений, обработку аудио и видео данных. Он может помочь в повышении качества данных, улучшении распознаваемости, улучшении передачи информации и восстановлении поврежденных данных.

II. МОДЕЛЬ ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Модель восстановления изображений на основе машинного обучения – это система, которая

использует алгоритмы машинного обучения для восстановления поврежденных или дефектных изображений и воспроизведения их оригинального внешнего вида.

Для восстановления изображения модель обучается на наборе данных, состоящем из искаженных и соответствующих оригинальных изображений. В процессе обучения модель изучает зависимости и структуру искаженных изображений и оригинальных изображений. Когда модель обучена, она может использоваться для восстановления новых, ранее не встречавшихся, поврежденных изображений.

Существуют разные подходы к моделированию восстановления изображений с использованием машинного обучения. Например, автоэнкодеры являются популярным выбором для моделирования восстановления изображений. Автоэнкодер это нейронная сеть, которая обучается сжимать и восстанавливать изображения. Во время обучения, автоэнкодер сжимает исходное изображение до некоторого скрытого представления, а затем восстанавливает его обратно к оригинальному изображению. При обучении модели подбираются оптимальные параметры для эффективного восстановления изображений.

Еще один подход – это использование генеративных моделей, таких как генеративные состязательные сети (GAN). GAN состоит из двух сетей: генератора, который пытается воссоздать оригинальное изображение из искаженного, и дискриминатора, который пытается различить между восстановленными и оригинальными изображениями. Обе сети учатся вместе, играя против друг друга, чтобы достичь наилучшего качества восстановленных изображений. Комбинированные модели, использующие как автоэнкодеры, так и генеративные модели, также могут быть использованы для восстановления изображений.

Модели восстановления изображений на основе машинного обучения имеют широкий спектр применений, включая восстановление старых или поврежденных фотографий, удаление шума или артефактов с изображений, а также увеличение разрешения изображения. Они могут быть полезны в таких областях, как медицинское образование, компьютерное зрение, изображение и видеобработка и многих других.

III. ПРОСТРАНСТВЕННАЯ ИНТЕРПОЛЯЦИЯ

Пространственная интерполяция – это метод, используемый для заполнения пропущенных значений данных в пространственной области. Он применяется в различных областях, таких как геоинформатика, компьютерное зрение и графика.

Пространственная интерполяция основана на анализе соседних значений и вычислении пропущенных значений, на основе имеющихся данных. Данные могут быть представлены в виде сетки или точечных образцов. Задача интерполяции состоит в том, чтобы восстановить пропущенные значения на основе известных значений.

Существует несколько методов пространственной интерполяции, применяемых в зависимости от особенностей данных и требуемых результатов. Некоторые из наиболее распространенных методов включают метод ближайших соседей, линейную интерполяцию, интерполяцию билинейной поверхностью, сплайн-интерполяцию и кригинг.

Метод ближайшего соседа заключается в присвоении пропущенным значениям ближайшего доступного значения. Этот метод прост и быстр, но может приводить к негладким результатам. Линейная интерполяция использует линейное соотношение между ближайшими известными значениями, чтобы вычислить пропущенные значения.

Интерполяция билинейной поверхностью основана на аппроксимации данных плоскостью, используя билинейную интерполяцию. Это дает более гладкие результаты, и он используется часто в графике и компьютерном зрении.

Сплайн-интерполяция использует кусочно-квадратичные или кубические функции для аппроксимации данных. Этот метод обеспечивает более гладкую интерполяцию и подходит для данных с кривыми или поверхностями.

Кригинг – это статистический метод интерполяции, который использует пространственные зависимости данных для предсказания пропущенных значений. Он основан на принципе, что близкие значения имеют большую связь, чем значения, находящиеся в дальнейшем удалении. Пространственная интерполяция является полезным инструментом для заполнения пропущенных значений и создания гладкой и непрерывной поверхности из предоставленных данных. Это помогает в анализе и визуализации данных, а также в различных приложениях обработки изображений и геоинформатики.

IV. ВОССТАНОВЛЕНИЕ ПО МЕТОДУ КРЫЛОВА

Восстановление по методу Крылова является численным методом решения систем линейных алгебраических уравнений.

Суть метода Крылова заключается в нахождении приближенного решения системы путем построения последовательности подпространств Крылова и нахождения наиболее близкого вектора из этой последовательности к искомому решению.

Первый шаг метода – выбор начального приближения для решения системы. Затем строится подпространство Крылова, которое формируется линейной комбинацией столбцов матрицы системы и векторов начального приближения.

Затем система уравнений разрешается по подпространству Крылова, что позволяет получить новое приближение к решению. После этого строится следующее подпространство Крылова и процесс повторяется до достижения требуемой точности решения.

Метод Крылова является эффективным для решения больших разреженных систем линейных уравнений, так как позволяет сократить количество операций. Однако его применение требует знания матричной структуры системы и может быть неустойчивым при наличии слишком больших или слишком малых собственных значений матрицы системы.

V. ЗАКЛЮЧЕНИЕ

Конкретный алгоритм восстановления изображений выбирается в зависимости от характера повреждений или искажений в изображении, доступного времени и доступных ресурсов для вычислений. Кроме того, важно принимать во внимание требования и ограничения конкретной задачи, доступность существующих реализаций и возможность адаптации алгоритмов под конкретные условия использования.

Иногда оптимальное решение может быть комбинацией различных алгоритмов восстановления, чтобы учесть разные аспекты и обеспечить наилучший результат.

СПИСОК ЛИТЕРАТУРЫ

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс // Техносфера – 2012. – С. 1105.
2. Яншин В., Калинин Г. Обработка изображений на языке СИ для IBM PC: Алгоритмы и программы / В. Яншин, Г. Калинин. // Мир – 1994. – С. 240.

АЛГОРИТМ ПОСТРОЕНИЯ ТРЁХМЕРНОЙ МОДЕЛИ ЧЕРЕПА ЧЕЛОВЕКА НА ОСНОВЕ АНАЛИЗА КТ-ИЗОБРАЖЕНИЙ

Курочка К. С., Карась О. В.
Кафедра информационных технологий,
Гомельский государственный технический университет имени П. О. Сухого
Гомель, Республика Беларусь
E-mail: olga.karas123@gmail.com

Данная статья посвящена вопросу построения трехмерных моделей черепа человека на основе анализа данных компьютерной томографии. В статье рассмотрены основные методы сегментации КТ-изображений и генерации 3D моделей – алгоритмы Marching Cubes и Dual Contouring. Приведена последовательность шагов обработки данных КТ, включающая предобработку, сегментацию, извлечение слоев и построение полигональной сетки методами Marching Cubes или Dual Contouring. Описаны типичные ошибки получаемых моделей и способы их исправления. В заключение сформулированы рекомендации по улучшению качества 3D реконструкции.

ВВЕДЕНИЕ

Построение трехмерных моделей черепа на основе КТ является актуальной задачей в различных областях медицины и биологии. 3D-реконструкция позволяет детально визуализировать анатомию пациента для диагностики патологий, планирования хирургического вмешательства, изучения биомеханики черепа. Кроме того, на основе индивидуальных моделей возможно изготовление имплантатов методом 3D-печати.

I. МЕТОДЫ ПОСТРОЕНИЯ ПОВЕРХНОСТИ ЧЕРЕПА ЧЕЛОВЕКА

Компьютерная томография позволяет получить поперечные срезы черепа, которые могут содержать не только изображения костей, но также и других тканей. Поэтому необходимо выполнить предварительную обработку КТ-изображений, включающую, как правило, фильтрацию и сегментацию данных с целью выявления костных структур [1, 2], формирующих поверхность черепа человека.

Для сегментации, на практике часто используют метод водораздела [3], метод среднего сдвига [4], GrabCut [5] и др. Это позволяет выделить костную ткань и определить контуры черепа на каждом слое, которые, впоследствии, будут использованы для создания трёхмерной модели.

Наиболее популярными методами трехмерной реконструкции объектов являются алгоритмы: марширующих кубов (Marching Cubes) [6] и двойных контуров (Dual Contouring) [7].

Алгоритм марширующих кубов используется в компьютерной графике для построения полигональной сетки по скалярному полю. Применяется для визуализации и обработки медицинских данных.

Работает путем перебора вокселей в объеме, определяя их положение на поверхности. Основные шаги:

- деление объема на сетку вокселей со скалярными значениями;
- классификация вокселей и определение полигональных шаблонов;
- вычисление пересечений ребер вокселей с поверхностью;
- генерация треугольников по пересечениям;
- объединение полигонов в общую сетку.

Алгоритм марширующих кубов имеет ограничения в виде "ступенчатости" сетки при низком разрешении. Можно оптимизировать установкой углов треугольников и распараллеливанием вычислений.

Алгоритм Dual Contouring позволяет строить гладкие и точные 3D модели на основе объемных данных. В отличие от Marching Cubes, он генерирует меньше полигонов на плоских участках и создаёт более плавные поверхности. Это достигается вычислением градиента функции в узлах сетки и оптимальным размещением вершин модели внутри ячеек. Алгоритм находит точки пересечения ребер ячейки с поверхностью, вычисляет градиенты функции в них и ищет оптимальное положение вершины ячейки. Основная трудность – выбор положения вершины при коллинеарных или противоположных градиентах. Несмотря на сложности, алгоритм дает высокое качество реконструкции и используется в современных 3D приложениях [7].

После построения 3D-модели необходимо убрать ошибки на модели, которые появляются из-за несовершенства используемого алгоритма Dual Contouring, а также из-за особенностей процедуры КТ-сканирования.

Возможные ошибки модели:

- self-intersection faces – обозначает наличие пересекающихся не по граням полигонов в модели;
- non-manifold vertices – обозначает наличие в модели вершин, которые используются более чем двумя полигонами;

- наличие лишних объектов рядом с моделью черепа, из-за наличия на изображениях КТ-элементов сканера, таких, как рамка фиксации головы пациента.

Для удаления всех вышеперечисленных проблем меша используется следующий алгоритм:

1. итеративно вызывается метод MeshAutoRepair библиотеки geometry3Sharp [8]. Количество итераций на этом шаге равно пяти;
2. из меша выбираются все соединённые между собой структуры полигонов. Таким образом получается несколько структур полигонов, не связанных между собой. Самой крупной структурой является череп, а все остальное – лишние объекты, которые необходимо удалить. Таким образом в меше остаётся только структура с наибольшим количеством полигонов, а всё остальное – игнорируется;
3. производится ремешинг модели по средствам метода FastestRemesh библиотеки geometry3Sharp. Данный метод позволяет сгладить острые грани модели, а также исправить большинство ошибок геометрии;
4. производится симплификация модели, для упрощения дальнейшей работы с программами слайсерами.

Конечным результатом является трехмерная модель поверхности человеческого черепа:

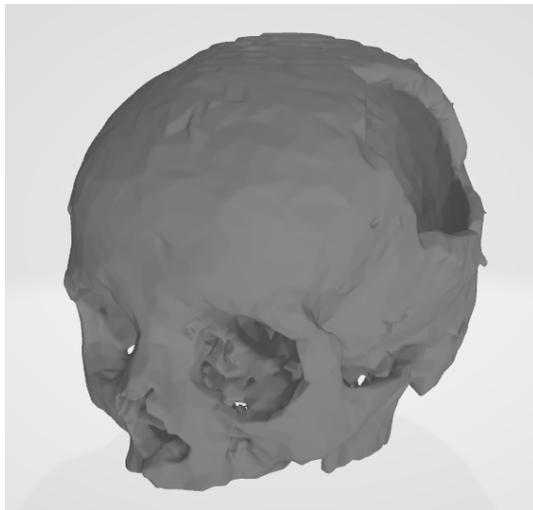


Рис. 1 – Поверхность черепа человека, построенная с помощью алгоритма Dual Contouring

Для улучшения результатов, полученных при использовании алгоритмов Marching Cubes и Dual Contouring, можно учесть следующие рекомендации:

- предварительно обрабатывать и улучшать входные данные для повышения точности;

- экспериментировать с различными пороговыми значениями для достижения желаемой детализации и точности;
- увеличивать разрешение данных;
- применять методы сглаживания (нормальная оценка, изменение положения вершин и лапласовское сглаживание) для уточнения поверхностей и удаления артефактов;
- применять адаптивные подходы, которые регулируют разрешение на основе особенностей локальных данных;
- оптимизировать реализацию для повышения производительности (параллельная обработка, использование памяти, аппаратное ускорение).

Следуя этим рекомендациям и учитывая конкретные требования приложения, можно повысить точность, качество и эффективность результатов, полученных с помощью алгоритмов Marching Cubes и Dual Contouring.

II. ЗАКЛЮЧЕНИЕ

Предлагаемая модификация алгоритма построения трёхмерной модели позволяет на основе DICOM файлов, содержащих результаты томографического исследования черепа человека, последовательно генерировать достаточно гладкие 3D-модели черепа, которые могут быть напечатаны на 3D-принтере.

СПИСОК ЛИТЕРАТУРЫ

1. Kurachka K. S., Tsalka I. M. Vertebrae detection in X-ray images based on deep convolutional neural networks // 2017 IEEE 14th International Scientific Conference on Informatics. – IEEE, 2017. – С. 194-196.
2. Kurochka K. S., Panarin K. A. An algorithm of segmentation of a human spine X-ray image with the help of Mask R-CNN neural network for the purpose of vertebrae localization // 2021 56th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST). – IEEE, 2021. – С. 55-58.
3. Прэнт, У. Цифровая обработка изображений / У. Прэнт. – М. Мир: 1982. – С. 502.
4. Машинное обучение: алгоритм кластеризации среднего сдвига [Электронный ресурс]. – Режим доступа: <https://russianblogs.com/article/66831200448/>. – Дата доступа: 09.10.2023.
5. Гонсалес, Р. С. Цифровая обработка изображений / Р. С. Гонсалес, Р. Е. Вудс. – Москва: 2012. – С. 186.
6. Golovanov, N. Geometric Modeling: The Mathematics of Shapes / N. Golovanov. – Natick, MA: A.K. Peters, 2002. – pp. 89.
7. Warren, J. Dual Contouring: A New Method for Generating 3D Hexahedral Meshes / J. Warren. – Proceedings of the Pacific Conference on Computer Graphics and Applications, 2004. – pp. 123.
8. DMesh3: A Dynamic Indexed Triangle Mesh [Электронный ресурс]. – Режим доступа: <http://www.gradient-space.com/tutorials/dmesh3>. – Дата доступа: 09.10.2023.

ФИЗИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ТРАНСКРАНИАЛЬНОЙ МАГНИТНОЙ СТИМУЛЯЦИИ

Мельникова В. В., Давыдов М. В.

Кафедра электроники, кафедра информатики,

Белорусский государственный университет информатики и радиоэлектроники

Объединённый институт проблем информатики Национальной академии наук Беларуси

Минск, Республика Беларусь

E-mail: melnikovav.v.work@gmail.com, davydov-mv@bsuir.by

На основе снимков МРТ была разработана методика построения гомогенной антропоморфной модели головного мозга. Была смоделирована трехмерная модель головы человека и различные виды индукторов в программном пакете SolidWorks. Затем проводился физический эксперимент с целью нахождения нужной концентрации и массовой доли солевого раствора для приближения электрической проводимости губчатого вещества к электрической проводимости серого вещества головного мозга. После этого, используя полученные в ходе эксперимента и программного моделирования данные, проводится эксперимент по физическому моделированию ТМС

ВВЕДЕНИЕ

В начале работы было сделано программное моделирование ТМС, которое можно разделить на несколько этапов:

1. Создание модели головы
2. Создание различных видов индукторов
3. Задание электрических параметров построенной модели и построение сетки
4. Результаты моделирования

Физический эксперимент также состоял из нескольких этапов:

1. Проведение исследования проводимости
2. Проведение ТМС на модели

I. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Метод создания модели головы. Построение модели выполнено на основе МРТ изображений, полученных из открытой базы данных OASIS [2]. Данные МРТ имеют расширение DICOM (*.DICOM). Импорт в программу 3DSlicer позволяет получить ряд изображений в трех проекциях.

На втором этапе выполнялась сегментация изображений. Сегментация вручную довольно долгий процесс, так как МРТ может содержать более 500 слоев для каждой проекции, 3D Slicer позволяет автоматизировать этот процесс. В результате использования встроенных функций каждый слой изображений будет разделен на 2 зоны.

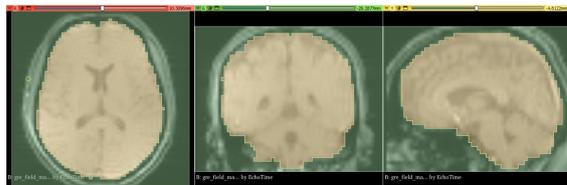


Рис. 1 – Сегментированные МРТ – изображения

Модель, полученная после преобразования имеет формат STL (*.stl), но для проведения

исследований в таких средах, как, например, COMSOL Multiphysics, эту модель следует конвертировать в твердое тело, для этого необходимо воспользоваться средой моделирования SOLIDWORKS.

Далее рассмотрим модели индукторов. При проведении ТМС используются основные виды индукторов: «большой кольцевой», «малый кольцевой», «восьмерка». Данные индукторы были созданы в среде моделирования SOLIDWORKS и затем импортированы для дальнейшего использования в COMSOL.

II. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

По результатам программного вычисления характеристик была проанализирована глубина и площадь воздействия каждого индуктора. Ниже приведен пример работы одного из индукторов, где с помощью цветной графики визуализировано воздействие на мозг индуктором «кольцевой большой».

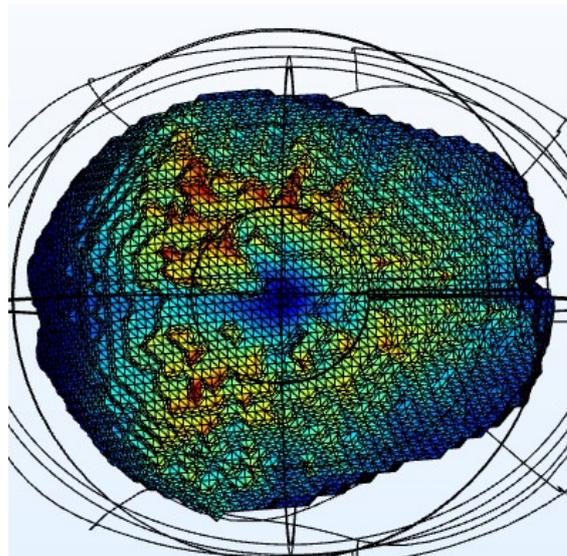


Рис. 2 – Плотность индукционного тока при ТМС гетерогенной антропоморфной модели мозга

При моделировании установлено, что использование различных типов индукторов позволяет изменять площадь стимуляции тканей мозга.

III. ФИЗИЧЕСКИЙ ЭКСПЕРИМЕНТ. ПРОВЕДЕНИЕ ИССЛЕДОВАНИЯ ПРОВОДИМОСТИ

Данное исследование проводилось с целью нахождения нужной концентрации и массовой доли солевого раствора для приближения электрической проводимости губчатого вещества к электрической проводимости серого вещества головного мозга. Подготовка к исследованию заключалась в создании электродов и закупке необходимых материалов: соль, губки, плата, медицинский шприц. При создании электродов использовалась плата, покрытая медью. Данная плата была разделена на две равные части (6x6 см). После разделения платы проводилось вытравливание меди методом кислотного травления (хлорид железа). Результат травления показан на рисунке 3.1. Расстояние между площадками составило 4 см, а размер самой площадки составляет 0.5x5 см.

Исследование проводилось при помощи прибора «Измеритель иммитанса E7-20». Губка, пропитанная солевым раствором, помещалась между двумя электродами, подключенными к измерителю иммитанса. Далее производилось снятие показаний.

Затем был сделан расчет электрической проводимости. Для расчетов использовалась формула удельного электрического сопротивления p :

$$p = \frac{R * S}{l}$$

где l – расстояние между электродами (4см), R – сопротивление, S – площадь поперечного сечения образца.

Так как электрическая проводимость является обратной величиной от удельного сопротивления, формула для расчета электрической про-

водимости g :

$$g = \frac{1}{p} = \frac{l}{R * S}$$

Анализ проводился с использованием научной публикации, из которой были взяты значения электрической проводимости серого вещества при разных частотах

Таблица 1 – Значения электрической проводимости серого вещества при разных частотах

Частота, Гц	Электрическая проводимость, С/м
1000	0,0988
4000	0,108
10000	0,115
100000	0,134

Был сделан вывод, что для получения электрической проводимости соответствующей серому веществу необходимо использовать дистиллированную воду массовой долей от 1 до 1,5 мл, либо 1% раствор соли с массовой долей около 0,5 мл. Стоит отметить, что полученные данные применимы исключительно для губчатого вещества размером 6x6 см, массой 1,8 г.

СПИСОК ЛИТЕРАТУРЫ

1. Транскраниальная магнитная стимуляция: клиническое применение и научные перспективы [Электронный ресурс]. – Режим доступа: <https://natural-sciences.ru/ru/article/view?id=35521>
2. Соловьева, Г. Р. Магнитотерапевтическая аппаратура / Г. Р. Соловьева. – М.: Медицина, 1991. – 176 с.
3. The Open Access Series of Imaging Studies. [Электронный ресурс]. Режим доступа: <https://www.oasis-brains.org>
4. COMSOL Multiphysics. [Электронный ресурс]. Режим доступа: <https://www.comsol.com>
5. Dielectric Properties » IT'IS Foundation [Электронный ресурс]. – Режим доступа: <https://itis.swiss/virtual-population/tissueproperties/database/dielectric-properties/>

АЛГОРИТМЫ ВЫЧИСЛЕНИЯ КОЛЛИЗИЙ В ВИРТУАЛЬНОМ ПРОСТРАНСТВЕ

Новиков П. А.

Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: zincore36@gmail.com

В статье проводится исследование и сравнительный анализ различных видов алгоритмов вычисления коллизий в виртуальном пространстве для оптимизации производительности и достижения точности в симуляции физического взаимодействия.

ВВЕДЕНИЕ

Вычисление коллизий между объектами в виртуальном пространстве является фундаментальной задачей в области компьютерных игр, виртуальной реальности, симуляторов в области медицины, автомобилестроения и робототехники, и других приложений, где требуется создание реалистичных и интерактивных виртуальных сцен. Точное и эффективное обнаружение столкновений между объектами является ключевым аспектом, влияющим на качество и производительность таких приложений.

На данный момент существует множество различных алгоритмов, наиболее эффективных в решении определенных задач: взаимодействие сложных объектов, множества простых объектов, обработка огромных или динамических сцен, симуляции жидкостей и мягких тел, и других.

1. ОСНОВНЫЕ ГРУППЫ АЛГОРИТМОВ

1. Алгоритмы разделения осей (Separating Axis Theorem, SAT).

Алгоритмы разделения осей (SAT) являются одними из наиболее распространенных алгоритмов для вычисления коллизий. Они основаны на принципе разделения осей и позволяют эффективно определить, пересекаются ли два выпуклых объекта. Алгоритм SAT проверяет наличие коллизий, проецируя объекты на различные оси и определяя, существует ли разделительный вектор между ними [1]. Если такой вектор существует, то объекты не пересекаются.

Преимущества:

- точность;
- возможность обработки сложных геометрических форм.

Недостатки:

- вычислительная затратность при большом количестве объектов;
- большое количество проверок коллизий.

2. Алгоритмы иерархий ограничивающих объемов (Bounding Volume Hierarchies, BVH).

Алгоритмы иерархий ограничивающих объемов (BVH) используются для оптимизации вычислений коллизий. BVH организует объекты в древовидную структуру, где каждый узел пред-

ставляет собой ограничивающий объем, такой как ограничивающий параллелепипед или сфера. Это позволяет эффективно отсеивать пары объектов, которые не могут столкнуться, уменьшая количество необходимых проверок коллизий.

В зависимости от типа деления ограничивающих объемов, существуют алгоритмы BVH с делением по середине (Midpoint Split), по площади (Surface Area Heuristic), по средней точке (Mean Split) и по объему (Volume Split).

Преимущества:

- высокая производительность;
- работа с большим количеством объектов;
- быстрое обнаружение простых коллизий.

Недостатки:

- низкая точность по сравнению с SAT;
- выделение дополнительных ресурсов для построения и обновления иерархии.

3. Алгоритмы сеток разбиения пространства (Spatial Partitioning Grid Algorithms).

Алгоритмы сеток разбиения пространства используются для упрощения вычислений коллизий путем разбиения виртуального пространства на сетку ячеек. Каждый объект помещается в соответствующую ячейку сетки, и затем проверяются только объекты, находящиеся в той же ячейке или смежных ячейках. Это позволяет сократить количество проверок коллизий, так как большинство объектов находятся в далеких ячейках и не могут пересекаться.

Преимущества:

- простота и эффективность для больших сцен с разреженными объектами;
- быстрое обнаружение коллизий между рядом находящимися объектами.

Недостатки:

- низкая точность при наличии сложных геометрических форм;
- дополнительные усилия для обработки объектов на пересечении ячеек сетки.

4. Геометрические алгоритмы.

Геометрические алгоритмы основаны на анализе геометрических характеристик объектов, таких как точки, линии и поверхности, для определения коллизий. Эти алгоритмы обычно требуют более сложных вычислений, но могут обеспечить

высокую точность при обнаружении и решении коллизий.

Преимущества:

- обработка сложных форм;
- более точные вычисления коллизий.

Недостатки:

- наиболее вычислительно затратные;
- требуют наибольшего количества ресурсов среди всех алгоритмов.

II. МЕТОДЫ СОКРАЩЕНИЯ ПРОВЕРКИ КОЛЛИЗИЙ (COLLISION CULLING TECHNIQUES)

Методы сокращения проверок коллизий используются для исключения ненужных проверок между парами объектов, которые могут быть гарантированно отсеяны без учета всех деталей геометрии. Некоторые из методов сокращения проверок коллизий включают:

1. Ограничивающие объемы (Bounding Volumes). Вместо проверки коллизий на основе полной геометрии объектов, используются ограничивающие объемы, такие как ограничивающие параллелепипеды или сферы, чтобы быстро определить, пересекаются ли объекты на основе их объемов. Виды ограничивающих объемов:

- Axis Aligned Bounding Boxes (AABB) – прямоугольный параллелепипед, который охватывает объект целиком, и определяется минимальным и максимальным значениями по каждой оси [1]. Применяет алгоритм Sweep and prune для нахождения столкновений [2].
- Bounding Spheres – сферические объемы, которые окружают объекты.
- Oriented Bounding Boxes (OBB) – ограничивающие параллелепипеды, имеющие произвольное направление и ориентацию в пространстве.

2. Временная согласованность (Time Coherence). Если движение объектов непрерывно, то можно использовать информацию о предыдущем состоянии объектов для прогнозирования и отсеивания некоторых пар объектов без дополнительной проверки коллизий [3]. Виды алгоритмов временной согласованности:

- Сохранение и восстановление (Save and Restore) – сохраняет промежуточные результаты вычислений на одном кадре и восстанавливает их на следующем кадре.
- Прогнозирование (Prediction) – предсказывает будущее состояние объектов на основе их предыдущего состояния и параметров движения.
- Кэширование (Caching) – кэширует результаты вычислений для объектов на предыдущих кадрах и повторно использует их на следующих кадрах.

- Упорядочение обновления (Update Ordering) – оптимизирует порядок обновления объектов или элементов сцены на последующих кадрах или временных шагах, на основе изменений и зависимостей между ними.

3. Пространственная разбивка (Spatial Partitioning). Использование методов разбиения пространства, таких как сетки разбиения пространства или иерархии ограничивающих объектов, позволяет эффективно отсеивать пары объектов, находящиеся далеко друг от друга [4]. Виды разбивки:

- Дерево октантов (Octree) – рекурсивно разделяет пространство на восемь равных кубических ячеек - октант.
- Сетка (Grid). Каждая ячейка сетки содержит информацию о объектах, находящихся внутри нее, что позволяет быстро определить, с какими объектами взаимодействует данная сущность.
- КД-дерево (KD-Tree) – разделяет пространство на оси-выделители (splitter axes) путем альтернативного разделения объектов по координатам X, Y и Z.
- R-дерево (R-Tree) – иерархическое дерево, которое разделяет пространство на ограничивающие прямоугольники. Каждый узел дерева содержит информацию о прямоугольниках и ссылки на поддеревья.

III. ЗАКЛЮЧЕНИЕ

Были рассмотрены основные алгоритмы вычисления коллизий объектов в виртуальном пространстве. Каждый из этих алгоритмов имеет свои преимущества и ограничения, и выбор конкретного алгоритма зависит от требований приложения и характеристик сцены.

Алгоритмы разделения осей (SAT) обеспечивают точность, но могут быть вычислительно затратными. Алгоритмы иерархий ограничивающих объемов (BVH) обеспечивают высокую производительность, но могут потерять в точности. Алгоритмы сеток разбиения пространства обеспечивают эффективность для разреженных сцен, но могут быть менее точными. Геометрические алгоритмы обеспечивают высокую точность, но требуют большего вычислительного ресурса.

СПИСОК ЛИТЕРАТУРЫ

1. Real-time Collision Detection. / Ericson, C. // Elsevier. – 2005. – № 1. – p. 632.
2. Collision detection between geometric models: a survey. / Ming C. Lin, Gottschalk S. // Proc. of IMA conference on mathematics of surfaces. – 1998. – Vol. 1. – pp. 602–608.
3. Physics for Game Developers. / Bourg M.D., Bywalec B. // O'Reilly Media. – 2013. – № 2. – p. 575.
4. Collision Detection in Interactive 3D Environments / Gino van den Bergen // CRC Press. – 2003. – № 1. – p. 277.

СПОСОБЫ ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ В ВИРТУАЛЬНЫХ ЧАСТНЫХ СЕТЯХ

Панасик А. А., Шилин Л. Ю., Хмыз Д. Д.

Кафедра вычислительных методов и программирования,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {a.panasik, shilin, d.khmyz}@bsuir.by

В данной статье рассматриваются виртуальные частные сети, выделены достоинства и недостатки виртуальных частных сетей, рассмотрены наиболее популярные VPN-сервисы.

ВВЕДЕНИЕ

Виртуальные частные сети (Virtual Private Networks, VPN) представляют собой технологию, которая играет важную роль в современной информационной безопасности и обеспечивает защиту конфиденциальности данных в сети Интернет. VPN обеспечивает безопасное и анонимное соединение между компьютером пользователя и удаленным сервером, что позволяет скрыть реальный IP-адрес и обеспечивает шифрование трафика. Эта технология имеет широкий спектр применений, начиная от обеспечения безопасного доступа к корпоративным ресурсам до обхода географических ограничений при доступе к контенту.

Работа VPN основана на принципе шифрования данных. Когда происходит подключение к VPN-серверу, все данные, передаваемые между устройством клиента и сервера, шифруются. Это делает практически невозможным перехват и просмотр данных третьими лицами, такими как хакеры, интернет-провайдеры или даже правительства.

Еще одним важным аспектом работы VPN является перенаправление трафика. Весь интернет-трафик идет через удаленный сервер, прежде чем попадать в сеть. Это дает возможность обходить цензуру и географические ограничения, так как можно выбрать VPN-сервер в любой стране, где доступ к определенным ресурсам не ограничен.

I. ПРИНЦИП РАБОТЫ VPN

Принцип работы VPN заключается в том, что все данные, передаваемые между пользователем и удаленным сервером, шифруются с использованием различных методов шифрования. Это обеспечивает конфиденциальность данных и защиту от несанкционированного доступа. Кроме того, VPN позволяют маскировать реальный IP-адрес пользователя, заменяя его на IP-адрес сервера, что делает его анонимным в сети.

1. Преимущества виртуальных частных сетей

1.1. Защита конфиденциальности и безопасность данных: VPN обеспечивает шифрование данных и скрытие IP-адреса пользователя, что

защищает информацию от перехвата и незаконного доступа. Это особенно важно, если предстоит работать с чувствительными данными или подключаться к открытым сетям Wi-Fi.

1.2. Безопасный удаленный доступ к корпоративным ресурсам: Для работы на удаленных проектах и доступа к внутренним серверам. VPN обеспечивает безопасное соединение с корпоративной сетью через интернет, что упрощает удаленную работу.

1.3. Обход цензуры и ограничений: В некоторых странах и сетях могут существовать ограничения на доступ к определенным веб-ресурсам. VPN позволяет обойти эти ограничения и получить доступ к всему мировому интернету.

1.4. Тестирование веб-сайтов и приложений: VPN может быть полезен для тестирования веб-сайтов и приложений с различных географических местоположений. Можно использовать серверы VPN в разных странах, чтобы проверить, как программный продукт работает в разных регионах.

1.5. Защита от атак и взломов: VPN может обеспечивать дополнительный уровень защиты от различных атак, таких как DDoS-атаки. Он может помочь скрыть настоящий IP-адрес, что делает атаку сложнее.

1.6. Анонимность и приватность: VPN позволяет скрывать физическое местоположение и делает вас более анонимными в сети. Это может быть полезно, если нужно оставаться анонимным при работе над проектами или исследованиями.

1.7. Безопасный обмен файлами: При обмене файлами между разработчиками, особенно в открытом интернете, VPN может усилить уровень безопасности, предотвращая возможные утечки данных.

1.8. Защита от слежки и мониторинга: VPN помогает снизить риск слежки и мониторинга онлайн-активности, что может быть важно для защиты личной и профессиональной конфиденциальности.

2. Недостатки виртуальных частных сетей

2.1. Уменьшение скорости передачи данных: Использование VPN может снизить скорость передачи данных из-за дополнительного шифрования и маршрутизации через удаленный сервер.

2.2. Надежность сервисов: Качество и надежность VPN-сервисов могут варьироваться. Не все провайдеры обеспечивают высокий уровень безопасности и конфиденциальности.

II. ИЗВЕСТНЫЕ И ШИРОКО ИСПОЛЬЗУЕМЫЕ VPN-СЕРВИСЫ

1. ExpressVPN – этот сервис известен своей высокой скоростью и широкой сетью серверов по всему миру.

2. NordVPN предлагает обширную сеть серверов и уделяет большое внимание безопасности.

3. CyberGhost – этот VPN-сервис славится своей простотой использования и обширной сетью серверов.

4. Hotspot Shield – VPN-сервис известен своим высоким уровнем безопасности и антивирусной защитой.

5. VyprVPN предлагает собственные серверы и акцентирует внимание на безопасности и конфиденциальности.

Преимущества и недостатки VPN-сервисов.

1. ExpressVPN. Преимущества: высокая скорость и стабильность соединения, сильное шифрование данных и защита от утечки DNS, обширная сеть серверов в разных странах, легкий в использовании клиент с интуитивным интерфейсом. Недостатки: относительно высокие цены по сравнению с некоторыми конкурентами.

2. NordVPN. Преимущества: безопасное и надежное соединение с высоким уровнем шифрования, огромная сеть серверов, включая специализированные серверы для различных целей, двойная VPN, обеспечивающая дополнительный уровень безопасности. Недостатки: скорость может быть несколько ниже, чем у некоторых конкурентов, не все серверы поддерживают P2P-трафик.

3. CyberGhost. Преимущества: простой в использовании клиент с различными опциями настройки, обширная сеть серверов., высокий уровень безопасности. Недостатки: скорость может быть переменной в зависимости от сервера и загрузки.

4. Hotspot Shield. Преимущества: защита от утечки DNS и высокий уровень безопасности, бесплатная версия доступна с ограничениями, отличная скорость. Недостатки: бесплатная версия содержит рекламу, премиум-подписка может быть дороже по сравнению с некоторыми конкурентами.

5. VyprVPN Преимущества: собственные серверы и высокий уровень безопасности, специализированные серверы для борьбы с цензурой и

обхода блокировок, хороший уровень скорости. Недостатки: ограничения на использование могут снизить удобство.

6. OpenVPN Преимущества: открытый исходный код, что позволяет более тщательно контролировать настройки и безопасность, может быть использован как самостоятельное решение или интегрировано в другие VPN-сервисы, высокая гибкость в настройке. Недостатки: требует технических навыков для настройки и использования, отсутствует интуитивный клиентский интерфейс.

III. ЗАКЛЮЧЕНИЕ

В заключение данной статьи о виртуальных частных сетях (VPN) следует подчеркнуть, что VPN являются важным инструментом в современном информационном мире. Они обеспечивают защиту данных, конфиденциальность и свободу доступа в сети. Программисты и многие другие профессионалы находят в VPN неоценимую помощь для обеспечения безопасности и эффективности своей работы.

VPN предоставляют возможность обходить цензуру и географические ограничения, защищать данные от несанкционированного доступа, а также обеспечивать безопасный доступ к корпоративным ресурсам. Они также способствуют защите личной конфиденциальности и анонимности в сети.

Однако при выборе VPN-сервиса важно тщательно оценить его характеристики, включая скорость, безопасность, цену и удобство использования. Каждый VPN-сервис имеет свои преимущества и недостатки, и правильный выбор зависит от индивидуальных потребностей пользователя.

В целом, VPN продолжают играть важную роль в сфере информационной безопасности, обеспечивая надежную защиту данных и сохранение конфиденциальности в онлайн-мире.

СПИСОК ЛИТЕРАТУРЫ

1. Раханов, К. Я. Обеспечение конфиденциальности информации в сети Интернет : пособие / К. Я. Раханов, Н. А. Раханова. – Новополюцк : Полоц. гос. ун-т, 2021. – 192 с.
2. Christine Johansen 10 Best VPN Services of 2023 – Top VPNs Tested By Experts / Christine Johansen [Электронный ресурс] // vpnmentor : [сайт]. – URL: <https://www.vpnmentor.com/> (дата обращения: 15.10.2023).
3. Choosing the VPN That's Right for You / [Электронный ресурс] // SURVEILLANCE SELF-DEFENSE : [сайт]. – URL: <https://ssd.eff.org/module/choosing-vpn-thats-right-you> (дата обращения: 16.10.2023).

АЛГОРИТМ ПОИСКА ПУТИ В ИГРОВОМ ПРИЛОЖЕНИИ. АЛГОРИТМ ДЕЙКСТРЫ

Рязанцев Д. Д., Жиляк Н. А.

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

Минск, Республика Беларусь

E-mail: {dmitryryaz, gh_nadya}@mail.ru

В статье описывается понятие алгоритма поиска пути. В частности рассматривается алгоритм Дейкстры

ВВЕДЕНИЕ

По мере развития игр увеличивались их размеры и сложность. Виртуальные миры и карты стали больше, как и количество взаимодействующих объектов. Нередко сотни объектов одновременно ищут путь, а время обработки, отведенное на поиск пути, существенно не увеличивается. В связи с этим разработчики часто вынуждены идти на компромиссы в алгоритмах поиска пути и тратить значительное время на настройку и проверку алгоритмов.

I. ОПРЕДЕЛЕНИЕ АЛГОРИТМА ПОИСКА ПУТИ

Алгоритм поиска пути в игровых приложениях – это алгоритм, который используется для нахождения пути между двумя точками на игровом поле. Он может быть использован для перемещения персонажей, NPC, машин и других объектов в игре. Существует множество алгоритмов поиска пути, каждый из которых имеет свои преимущества и недостатки.

II. РАЗНОВИДНОСТИ АЛГОРИТМОВ ПОИСКА ПУТИ

Среди наиболее распространенных алгоритмов поиска пути в играх можно выделить:

1. Алгоритм Дейкстры: Этот алгоритм используется для нахождения кратчайшего пути между двумя точками на графе. Он может быть использован для поиска пути в играх, где перемещение ограничено определенными направлениями или где есть препятствия на пути.
2. алгоритм поиска в ширину: Этот алгоритм используется для нахождения кратчайшего пути между двумя точками на графе. Он может быть использован для поиска пути в играх, где перемещение не ограничено определенными направлениями;
3. алгоритм поиска в глубину: Этот алгоритм используется для нахождения всех возможных путей между двумя точками на графе. Он может быть использован для поиска путей в играх, где перемещение не ограничено определенными направлениями;

4. алгоритм Ли: Этот алгоритм используется для нахождения кратчайшего пути между двумя точками на графе с единичными стоимостями ребер. Он может быть использован для поиска путей в играх, где перемещение не ограничено определенными направлениями;

5. алгоритм Флойда-Уоршелла: Этот алгоритм используется для нахождения кратчайших расстояний между всеми парами вершин в графе. Он может быть использован для поиска кратчайших расстояний между всеми объектами в игре.

Во всех этих алгоритмах соблюдается баланс между временем поиска и используемой памятью

Выбор наиболее подходящего алгоритма зависит от среды поиска пути в видеоигре и ее требований.

Существуют три свойства, которыми должны обладать алгоритмы, чтобы быть полезными на практике:

- Согласованность решений - качество решений (субоптимальность) не должно резко отличаться в разных задачах;
- адекватное время отклика - жесткие ограничения на время отклика, диктуемые игрой, должны быть соблюдены;
- Эффективность использования памяти - объем памяти для одного объекта и фиксированной памяти должен быть сведен к минимуму.

Цель - минимизировать неоптимальность, время отклика и использование памяти.

III. АЛГОРИТМ ДЕЙКСТРЫ И ЕГО ОСОБЕННОСТИ

Алгоритм Дейкстры – это алгоритм на графах, который используется для нахождения кратчайшего пути между двумя точками. Он может быть использован для поиска пути в играх, где перемещение ограничено определенными направлениями или где есть препятствия на пути. Алгоритм Дейкстры работает только для графов без ребер отрицательного веса.

Алгоритм Дейкстры начинает работу с одной из вершин графа и находит кратчайшие пути от этой вершины до всех остальных. Он работает

пошагово, на каждом шаге он “посещает” одну вершину и пытается уменьшать метки. Работа алгоритма завершается, когда все вершины посещены.

Примеры игр, в которых используется алгоритм Дейкстры: Civilization, Starcraft, World of Warcraft, Diablo, Half-Life 2.

Пример процесса вычисления алгоритма Дейкстры:

Если имеется четыре узла 1, 2, 3, 4, с заданными узлами, линиями и весами (см. рис. 1):

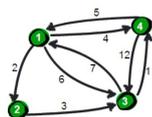


Рис. 1 – Пример работы алгоритма Дейкстры

Тогда в результате вычисления кратчайшего пути от 1 до всех точек с помощью алгоритма Дейкстры получим следующие пути: $\{1 \rightarrow 2\}$ $\{1 \rightarrow 4\}$ $\{1 \rightarrow 2 \rightarrow 3\}$.

На первой итерации, начиная непосредственно с узла 1, находим узел, на который он указывает, и заполняем соответствующую ячейку с весом до узла. Если путь к точке невозможен, соответствующая графа заполняется бесконечностью. Затем сравниваем веса каждого узла на первой итерации, и кратчайший путь от узла 1 к узлу 2 равен $\{1 \rightarrow 2\}$, а вес равен 2.

Во второй итерации на основе предыдущей итерации заполнить бесконечность в ячейках, соответствующих путей. Затем расходятся по текущим покрытым узлам для поиска кратчайшего пути. Выясняется, что вес пути из $\{1 \rightarrow 2 \rightarrow 3\}$ равен 5, что меньше исходного значения веса непосредственно из $1 \rightarrow 2$, поэтому значение веса записывается в ячейку, соответствующую 3. Наконец, сравниваются значения веса каждого узла, оставшегося на второй итерации, и кратчайший путь из узла 1 в узел 4 получается как $\{1 \rightarrow 4\}$, а вес равен 4.

В третьей итерации, поскольку в графе осталось только четыре узла, эта итерация является последней. На основании предыдущей, покрытые в данный момент узлы расходятся в окрестности для поиска более короткого пути. Однако более короткого пути найти не удастся, и остаются только узлы, которые можно выбрать, поэтому кратчайший путь из узла 1 в узел 3 получается, как $\{1 \rightarrow 2 \rightarrow 3\}$, а вес равен 5.

Наконец, получен кратчайший путь от узла 1 к каждому узлу.

IV. ИДЕИ ПРИМЕНЕНИЯ АЛГОРИТМА ДЕЙКСТРЫ В ИГРОВЫХ ПРИЛОЖЕНИЯХ

Конкретные идеи применения алгоритма Дейкстры для интеллектуального поиска пути в игре таковы:

1. Во-первых, на карту должны быть добавлены виртуальные горизонтальные и вертикальные пересекающиеся прямые линии, чтобы создать спецификационную сеть, похожую на шашечную сетку, тогда пересечение двух прямых линий или центр шашечной доски будет использовано в качестве узла. При этом ИИ может выполнять только четыре смежные операции - движение вверх, вниз, влево и вправо, поэтому матрица смежности графа задается только в четырех смежных направлениях каждого узла, причем те, которые выходят за пределы диапазона количества узлов, не задаются, а остальные узлы задаются до бесконечности;
2. Поскольку ИИ также необходим для управления персонажем, чтобы завершить движение по орбите тела столкновения, строки и столбцы матрицы смежности, в которых находится тело столкновения, все устанавливаются на бесконечность, то есть линия, соединяющая тело столкновения, разрывается.

Недостатки:

- Если масштаб сети слишком велик, то получение результата будет занимать много времени для получения результата;
- Для некоторых приложений, чувствительных к времени, или сервисов, работающих в реальном времени, нам необходимо сократить время работы.

V. ЗАКЛЮЧЕНИЕ

В данной работе были рассмотрены алгоритмы поиска пути, в частности алгоритм Дейкстры, который является одним из наиболее широко используемых алгоритмов для поиска кратчайшего пути в графе. Он обеспечивает точные и эффективные результаты, что делает его идеальным выбором для использования в игровых приложениях. Однако стоит отметить, что выбор алгоритма поиска пути зависит от конкретных требований и ограничений игры. В некоторых случаях другие алгоритмы, такие как A^* или BFS, могут быть более подходящими. В любом случае, правильное понимание и применение этих алгоритмов является ключом к созданию успешных и захватывающих игровых приложений.

СПИСОК ЛИТЕРАТУРЫ

1. Fast Grid-based Path Finding for Video Games / William Lee, Ramon Lawrence // University of British Columbia
2. Study of the Interception Scheme Based on A^* Path Finding Algorithm in Computer Game / Zhenjie Li, Haiming Li // Journal of Computer and Communications, 2020, 8, 32-49
3. ProgrammerSought [Электронный ресурс] /Application of Dijkstra Algorithm in Game Intelligent Pathfinding - Режим доступа: [urlhttps://www.programmersought.com/article/23446213813/](https://www.programmersought.com/article/23446213813/) - Дата досутпа: 13.10.2023

АЛГОРИТМ РАСПОЗНАВАНИЯ ДВИЖЕНИЙ ЧЕЛОВЕКА

Рязанцев Н. Д., Жиляк Н. А.

Кафедра вычислительных методов и программирования,
Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {ryazantsev.nikitka01, gh_nadya}@mail.ru

В статье описывается понятие алгоритма распознавания движений человека. В частности рассматривается Алгоритм HOG (Histogram of Oriented Gradients)

ВВЕДЕНИЕ

Распознавание движений человека – это процесс, который позволяет компьютеру распознавать и классифицировать движения человека. Эта технология нашла широкое применение в различных областях, таких как медицина, спорт, игровая индустрия и другие. С развитием технологий распознавания движений человека появилась возможность создания новых приложений и устройств, которые могут быть использованы в различных областях. Например, в медицине это может быть использовано для диагностики заболеваний, в спорте – для анализа движений спортсменов, а в игровой индустрии – для создания новых игр.

I. ПОДХОДЫ К РАСПОЗНАВАНИЮ ДВИЖЕНИЙ ЧЕЛОВЕКА

Существует множество подходов к распознаванию движений человека, включая методы, основанные на глубоком обучении, методы, основанные на машинном обучении, и методы, основанные на компьютерном зрении. Некоторые из этих методов используют данные с акселерометров и гироскопов, а другие – данные с камер.

Одним из наиболее популярных методов распознавания движений человека является метод, основанный на глубоком обучении. Этот метод использует нейронные сети для обработки данных и выявления закономерностей в движениях человека. Однако этот метод требует большого количества данных для обучения и может быть сложен в реализации.

Другой популярный метод – это метод, основанный на компьютерном зрении. Этот метод использует камеры для записи движений человека и анализа этих записей для выявления закономерностей в движениях. Этот метод может быть более точным, чем методы, основанные на глубоком обучении, но может быть менее эффективным в условиях ограниченного пространства или освещения.

Наконец, существуют методы, основанные на машинном обучении. Эти методы используют алгоритмы машинного обучения для выявления закономерностей в данных о движениях человека. Эти методы могут быть менее точными, чем

методы, основанные на глубоком обучении или компьютерном зрении, но могут быть более эффективными в условиях ограниченных ресурсов.

Алгоритм HOG может быть использован для распознавания объектов на изображениях, таких как лица, автомобили и пешеходы. Он также может быть использован для распознавания жестов и других движений человека. Алгоритм HOG широко используется в различных областях, таких как медицина, безопасность и игровая индустрия.

II. АЛГОРИТМ HOG (HISTOGRAM OF ORIENTED GRADIENTS)

Алгоритм HOG (Histogram of Oriented Gradients) был впервые описан в 2005 году исследователями INRIA Навнитом Далалом и Биллом Триггсом. В своей работе они использовали алгоритм для нахождения пешеходов на статичных изображениях, хотя впоследствии расширили область применения до нахождения людей на видео, а также различных животных и машин на статичных изображениях. Алгоритм HOG основан на подсчете количества направлений градиента в локальных областях изображения. Этот метод похож на гистограммы направления края, дескрипторы SIFT и контексты формы, но отличается тем, что вычисляется на плотной сетке равномерно распределенных ячеек и использует нормализацию перекрывающегося локального контраста для увеличения точности. Алгоритм HOG работает следующим образом:

1. Вычисление градиента: Изображение разбивается на маленькие связные области, называемые ячейками. Для каждой ячейки вычисляются градиенты интенсивности пикселей;
2. Группировка направлений: Направления градиентов объединяются в гистограммы направленных градиентов;
3. Блоки дескрипторов: Ячейки объединяются в блоки, и для каждого блока вычисляется дескриптор, который представляет собой нормализованную версию гистограмм направленных градиентов;
4. Нормализация блоков: Дескрипторы блоков нормализуются по контрасту, чтобы уменьшить влияние изменений освещения на результаты распознавания.

В алгоритме НОГ используются два основных метода нормализации гистограмм:

1. Глобальная нормализация изображения: Этот этап является необязательным и предназначен для уменьшения влияния эффектов освещения. На практике используется гамма-компрессия (закон степени), которая вычисляет квадратный корень или логарифм каждого цветового канала. Сила текстуры изображения обычно пропорциональна местному освещению поверхности, поэтому такая компрессия помогает уменьшить влияние локального затенения и изменений освещения;
2. Нормализация блоков: Этот этап направлен на создание кодировки, которая чувствительна к локальному содержанию изображения, оставаясь при этом устойчивой к небольшим изменениям позы или внешнего вида. Метод заключается в накоплении меры локальной энергии гистограммы по локальным группам ячеек, которые называются «блоками». Полученный результат используется для нормализации каждой ячейки в блоке. Обычно каждая отдельная ячейка совместно используется несколькими блоками, но ее нормализации зависят от блока и, следовательно, различны.

Эти методы нормализации обеспечивают устойчивость алгоритма НОГ к изменениям освещения и теней, что делает его эффективным инструментом для обнаружения объектов на изображениях.

Алгоритм НОГ может быть использован для распознавания объектов на изображениях, таких как лица, автомобили и пешеходы. Он также может быть использован для распознавания жестов и других движений человека. Алгоритм НОГ широко используется в различных областях, таких как медицина, безопасность и игровая индустрия.

III. СПОСОБЫ ОПТИМИЗАЦИИ АЛГОРИТМА

Существуют различные методы оптимизации алгоритма НОГ, которые можно классифицировать в три группы:

1. Методы, которые улучшают определенный этап алгоритма: Эти методы фокусируются на оптимизации конкретных этапов алгоритма НОГ, например, вычисления градиента или нормализации гистограммы;
2. Методы, которые оптимизируют весь алгоритм: Эти методы включают в себя изменения во всем алгоритме с целью улучшения его производительности. Это может включать в себя изменение способа вычисления

гистограмм, использование различных методов нормализации или изменение структуры ячеек и блоков;

3. Методы, которые делают незначительное упрощение алгоритма: Эти методы включают в себя небольшие упрощения алгоритма, которые могут привести к увеличению скорости выполнения без значительного снижения точности.

Алгоритм НОГ может быть затратен по времени и ресурсам, особенно при работе с большим количеством изображений или видео. Однако, существуют методы ускорения работы алгоритма НОГ, такие как использование каскадных классификаторов.

Каскадный классификатор – это алгоритм, который используется для быстрого и эффективного поиска объектов на изображении. Он работает путем использования нескольких слабых классификаторов, каждый из которых проверяет изображение на наличие объекта. Если слабый классификатор обнаруживает объект, он передает изображение следующему слабому классификатору. Если объект не обнаружен, изображение отбрасывается. Таким образом, каскадный классификатор может быстро отбросить большое количество изображений, которые не содержат объекты, что ускоряет работу алгоритма НОГ.

IV. ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ АЛГОРИТМА

в 2019 году исследователи из университета Карнеги-Меллона использовали алгоритм НОГ для распознавания жестов рук. Они провели эксперименты на датасете, состоящем из 27 жестов рук, и получили точность распознавания более 90 %. В другом исследовании, проведенном в 2018 году, алгоритм НОГ был использован для распознавания движений человека в видео. Исследователи использовали алгоритм НОГ для распознавания движений человека на видео, полученном с камеры наблюдения, и получили точность распознавания более 80 %.

СПИСОК ЛИТЕРАТУРЫ

1. A Revisit Histogram of Oriented Descriptor for Facial Color Image Classification Based on Fusion of Color Information / Huy Nguyen-Quoc, Vinh Truong Hoang // Ho Chi Minh City Open University
2. An optimal method based on HOG-SVM for fault detection / Panfeng Xu, Lidong Huang, Yan Song // Liaoning University
3. Iq.opengenus [Электронный ресурс] /Using Histogram of Oriented Gradients (HOG) for Object Detection – Режим доступа: <https://iq.opengenus.org/object-detection-with-histogram-of-oriented-gradients-hog/> – Дата доступа: 13.10.2023

АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧИ УПРАВЛЕНИЯ СКЛАДСКИМИ ЗАПАСАМИ

Устинович К. А., Лабоцкий В. В.
Кафедра управления информационными ресурсами,
Академия управления при Президенте Республики Беларусь
Минск, Республика Беларусь
E-mail: kseniya.ustinovich.01@gmail.com

Снижение объема пассажирских перевозок порождает у транспортных предприятий дефицит финансовых источников повышения конкурентоспособности оказываемых услуг (своевременного и эффективного обновления подвижного состава, обеспечения регламентированных показателей выхода подвижного состава на линию). Низкий уровень конкурентоспособности, в свою очередь, приводит к неудовлетворенности потребителей качеством услуг и поиску иных способов удовлетворения своих потребностей в перемещении в пространстве за счет приобретения личного автотранспорта или частичного отказа от поездок на общественном. Таким образом, возникает последовательная взаимосвязь явлений и событий, которые приводят к образованию целого комплекса системных проблем в функционировании автомобильного пассажирского транспорта.

I. МЕТОДЫ И АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧ УПРАВЛЕНИЯ СКЛАДСКИМИ ЗАПАСАМИ

Для организации бесперебойной работы пассажирского автомобильного транспорта рекомендуется внедрение автоматизированной модели управления запасами. Модель управления запасами включает: выбор и обоснование критерия оптимизации (чаще всего это минимум логистических издержек, связанных с управлением запасами, может быть минимальное время выполнения заказа, максимальная надежность поставки), расчет издержек управления запасами, формулировку ограничений, моделирование спроса (расхода) и пополнения запасов, расчет стратегии управления. Выделяют 2 основные модели управления запасами:

1. Модель с фиксированным размером заказа.
2. Модель с фиксированным интервалом времени между заказами.

А также существует 2 дополнительные модели управления запасами: Выделяют 2 основные модели управления запасами:

1. Модель с установленной периодичностью пополнения запасов до установленного уровня;
2. Модель «максимум-минимум».

Решение задачи управления складскими запасами требует применения различных алгоритмов и подходов, таких как:

1. Алгоритм прогнозирования спроса: предсказание будущего спроса на товары. Он может использовать статистические методы, такие как методы временных рядов (например, скользящее среднее, экспоненциальное сглаживание) или методы регрессии для анализа исторических данных о спросе и создания прогнозов.
2. Алгоритм определения уровня безопасности: помогает определить оптимальный уровень запасов, при котором не возникнет де-

фицита. Он может использовать методы, такие как анализ потребления, статистический анализ вариаций, идентификацию точек сбора заказов и расчет вероятности дефицита.

3. Алгоритм расчета точек заказа: определяет, когда нужно делать заказ на пополнение запасов. Он учитывает уровень безопасности, время доставки товара и потребность в товаре. Этот алгоритм может использовать параметры, такие как время доставки, затраты на хранение запасов и экономическую партию заказа.
4. Алгоритмы оптимизации запасов: подбирают оптимальные значения для параметров управления запасами, такие как количество заказываемого товара и уровень безопасности. Они могут использовать методы, такие как статическая и динамическая оптимизация запасов, методы линейного программирования и эвристические алгоритмы поиска оптимальных решений.
5. Алгоритмы классификации и сегментации: помогают классифицировать товары по их значимости, спросу, скорости оборачиваемости и другим параметрам. Они могут использовать методы, такие как анализ ABC/XYZ, кластерный анализ, ассортиментный анализ и прогнозирование оборачиваемости.

Это лишь некоторые из алгоритмов, используемых в задаче управления складскими запасами. Реальные модели управления запасами комбинируют несколько алгоритмов и адаптируются к конкретным бизнес-процессам и требованиям организации.

II. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ РЕШЕНИЯ ЗАДАЧ УПРАВЛЕНИЯ СКЛАДСКИМИ ЗАПАСАМИ

Для сокращения трудоемкости и временных затрат разработано программное средство «Моде-

ли управления запасами» на языке Visual Basic. Разработанное программное приложение существенно облегчает процесс планирования пополнения запасов специалистами предприятий в виду простоты интерфейса. При этом рутинная работа выполняется в автоматическом режиме увеличивая эффективность использования рабочего времени. Эффективность работы данного приложения может повыситься путем считывания исходных данных с имеющихся таблиц, которые составляются как анализ на предприятии. Также возможно сохранение полученных данных в виде рассчитанных значений и построенных графиков. Вводятся данные через TextBox: ежемесячное потребление, время выполнения заказа, объем потребления, затраты на выполнение одного заказа, затраты на содержание единицы запаса, время задержки поставки. Эта информация либо в полученном виде либо в предварительно обработанном виде помещается в Лист Excel (см. рис. 1).

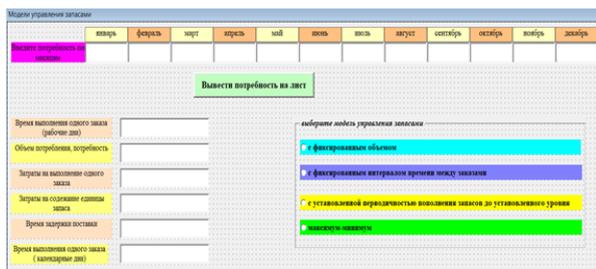


Рис. 1 – Вид рабочего окна программного средства

Для всех моделей управления запасами, которые расположены в OptionButton, активируем Лист Excel для вывода данных и очищаем диапазон ячеек, а также деактивируем объект VBA Chart для ранее созданных отчетов, который используется для построения диаграммы. Для построения более точной модели был написан участок кода, который отслеживает количество дней в каждом календарном месяце. Такая реализация была выбрана для того, чтобы получить более корректные результаты с учетом среднего днев-

ного потребления и увидеть в какой именно день запас был в дефиците или пополнился. По рассчитанным данным строится график движения запасов с помощью объекта VBA Chart.

III. Выводы

Модель с установленной периодичностью пополнения запасов до установленного уровня является оптимальной для пассажирских автотранспортных предприятий в связи с тем, что она предоставляет повышенную защищенность от дефицита запасов. Чтобы предотвратить завышение объемов запасов, содержащихся на складе, или их дефицит, заказы производятся не только в установленные моменты времени, но и при достижении запасом порогового уровня. Таким образом, рассматриваемая система включает в себя элемент системы с фиксированным интервалом времени между заказами (установленную периодичность оформления заказа) и элемент системы с фиксированным размером заказа (отслеживание порогового уровня запасов).

СПИСОК ЛИТЕРАТУРЫ

1. Алексева, Н. И. Управление товарными запасами предприятия на основе анализа системы показателей / Н. И. Алексева, Ю. В. Алехина // Стратегия предприятия в контексте повышения его конкурентоспособности – 2019. – Т. 1, № 5(5). – с. 56–59.
2. Данилова, С. Ю. Оптимизация уровня запасов производства с целью повышения конкурентоспособности предприятия с непрерывным циклом производства / С. Ю. Данилова, Е. В. Пуденков // Вестник Волжского университета имени В. Н. Татищева – 2019. – № 2 (28). – с. 47–52.
3. Маклаков, Г. В. Эффективность коммерческой деятельности / Г. В. Маклаков // Издательство: М. Дашков и К – 2020. – 232 с.
4. Сухотерин, Д. А. Эффективное управление складскими запасами как фактор повышения рентабельности оборотных средств / Д. А. Сухотерин // Финансовые рынки и банки – 2021. – Режим доступа: <https://cyberleninka.ru/article/n/effektivnoe-upravlenie-skladskimi-zapasami-kak-faktor-povysheniya-rentabelnosti-oborotnyh-sredstv-predpriyatiya>. – Дата доступа: 18.10.2023.

АНАЛИТИЧЕСКИЕ СВОЙСТВА РЕШЕНИЙ СЕМЕЙСТВА ТРЕХМЕРНЫХ ДИНАМИЧЕСКИХ ПЯТИЭЛЕМЕНТНЫХ СИСТЕМ С ОДНОЙ КВАДРАТИЧНОЙ НЕЛИНЕЙНОСТЬЮ И ЧЕТЫРЬМЯ ЛИНЕЙНЫМИ ЭЛЕМЕНТАМИ

Цегельник В. В.

Кафедра высшей математики,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: tsegvv@bsuir.by

Исследован характер возможных подвижных особых точек решений семейства трехмерных пятиэлементных динамических систем с одной квадратичной нелинейностью и четырьмя линейными элементами.

ВВЕДЕНИЕ

В работе [1] выполнено качественное исследование решений трехмерных пятиэлементных систем с одной квадратичной нелинейностью с произвольным параметром A и параметром $\varepsilon^2 = 1$. В системах указанного семейства параметр A входит как множитель при линейном элементе, либо как отдельный элемент (константа).

I. ПОСТАНОВКА ЗАДАЧИ

Целью настоящей работы является исследование характера подвижных особых точек (т.е. точек, положение которых зависит от начальных условий) решений систем [1] с одной квадратичной нелинейностью и четырьмя линейными элементами

$$\begin{aligned} \dot{x} &= y^2 + \varepsilon x, \dot{y} = x + Az, \dot{z} = x. & (12) \\ \dot{x} &= y^2 + \varepsilon x, \dot{y} = x + Az, \dot{z} = y. & (13) \\ \dot{x} &= y^2 + Ax, \dot{y} = x + z, \dot{z} = \varepsilon z. & (14) \\ \dot{x} &= y^2 + \varepsilon x, \dot{y} = Ay + z, \dot{z} = x. & (15) \\ \dot{x} &= y^2 + Ay, \dot{y} = x + z, \dot{z} = \varepsilon z. & (16) \\ \dot{x} &= y^2 + Ay, \dot{y} = \varepsilon y + z, \dot{z} = x. & (17) \\ \dot{x} &= y^2 + Ay, \dot{y} = x + \varepsilon y, \dot{z} = x. & (18) \\ \dot{x} &= y^2 + Az, \dot{y} = x + \varepsilon y, \dot{z} = y. & (19) \\ \dot{x} &= y^2 + z, \dot{y} = x + Ay, \dot{z} = \varepsilon z. & (20) \\ \dot{x} &= y^2 + Az, \dot{y} = x + z, \dot{z} = \varepsilon z. & (21) \\ \dot{x} &= y^2 + Az, \dot{y} = \varepsilon y + z, \dot{z} = x. & (22) \\ \dot{x} &= z^2 + Ax, \dot{y} = x + \varepsilon y, \dot{z} = y. & (23) \\ \dot{x} &= z^2 + \varepsilon x, \dot{y} = x + Az, \dot{z} = y. & (24) \\ \dot{x} &= z^2 + Ay, \dot{y} = x + \varepsilon y, \dot{z} = x. & (25) \\ \dot{x} &= z^2 + Ay, \dot{y} = x + \varepsilon y, \dot{z} = y. & (26) \\ \dot{x} &= z^2 + Ay, \dot{y} = \varepsilon y + z, \dot{z} = x. & (27) \\ \dot{x} &= z^2 + Az, \dot{y} = x + \varepsilon y, \dot{z} = y. & (28) \\ \dot{x} &= yz + Ax, \dot{y} = x + \varepsilon y, \dot{z} = x. & (29) \\ \dot{x} &= yz + \varepsilon x, \dot{y} = x + Az, \dot{z} = x. & (30) \end{aligned}$$

$$\begin{aligned} \dot{x} &= y^2 + \varepsilon x + Az, \dot{y} = x, \dot{z} = y. & (1) \\ \dot{x} &= y^2 + Ax + z, \dot{y} = x, \dot{z} = \varepsilon z. & (2) \\ \dot{x} &= y^2 + Ay + z, \dot{y} = x, \dot{z} = \varepsilon z. & (3) \\ \dot{x} &= y^2 + \varepsilon x + Ay, \dot{y} = z, \dot{z} = x. & (4) \\ \dot{x} &= y^2 + \varepsilon x + Az, \dot{y} = z, \dot{z} = x. & (5) \\ \dot{x} &= yz + \varepsilon x + Ay, \dot{y} = z, \dot{z} = x. & (6) \\ \dot{x} &= \varepsilon x + y + Az, \dot{y} = x^2, \dot{z} = x. & (7) \\ \dot{x} &= \varepsilon x + y + Az, \dot{y} = x^2, \dot{z} = y. & (8) \\ \dot{x} &= Ax + y + z, \dot{y} = x^2, \dot{z} = \varepsilon z. & (9) \\ \dot{x} &= \varepsilon x + y + Az, \dot{y} = z^2, \dot{z} = x. & (10) \\ \dot{x} &= \varepsilon x + y + Az, \dot{y} = xz, \dot{z} = y. & (11) \end{aligned}$$

$$\dot{x} = yz + \varepsilon x, \dot{y} = x + Az, \dot{z} = y. \quad (31)$$

$$\dot{x} = yz + Ax, \dot{y} = \varepsilon y + z, \dot{z} = x. \quad (32)$$

$$\dot{x} = yz + Ay, \dot{y} = x + \varepsilon y, \dot{z} = x. \quad (33)$$

$$\dot{x} = yz + Ay, \dot{y} = \varepsilon y + z, \dot{z} = x. \quad (34)$$

$$\dot{x} = yz + Ay, \dot{y} = x + \varepsilon y, \dot{z} = x. \quad (35)$$

$$\dot{x} = yz + Az, \dot{y} = x + \varepsilon y, \dot{z} = y. \quad (36)$$

$$\dot{x} = yz + Az, \dot{y} = \varepsilon y + z, \dot{z} = x. \quad (37)$$

$$\dot{x} = \varepsilon x + Ay, \dot{y} = x + z, \dot{z} = x^2. \quad (38)$$

$$\dot{x} = \varepsilon x + y, \dot{y} = Ay + z, \dot{z} = x^2. \quad (39)$$

$$\dot{x} = \varepsilon x + Ay, \dot{y} = x + z, \dot{z} = y^2. \quad (40)$$

$$\dot{x} = \varepsilon x + Az, \dot{y} = x + z, \dot{z} = y^2. \quad (41)$$

$$\dot{x} = \varepsilon x + Ay, \dot{y} = x + z, \dot{z} = xy. \quad (42)$$

$$\dot{x} = \varepsilon x + y, \dot{y} = Ay + z, \dot{z} = xy. \quad (43)$$

$$\dot{x} = \varepsilon x + z, \dot{y} = Ay + z, \dot{z} = xy. \quad (44)$$

с неизвестными функциями x, y, z в предположении, что неизвестная переменная t является комплексной; $\varepsilon^2 = 1$, A – произвольный постоянный параметр.

Системы (2), (9), (14), (15), (20), (23), (29), (32), (39), (43), (44) являются диссипативными, если $A + \varepsilon < 0$. Остальные системы из данного списка являются таковыми, если $\varepsilon = -1$.

Системы (уравнения), общие решения которых не имеют подвижных критических особых точек называют системами (уравнениями) Пенлеве-типа или P -типа.

II. АЛГОРИТМ

Для решения поставленной задачи использован тест Пенлеве [2], представляющий набор условий, необходимых для отсутствия у общего решения системы дифференциальных уравнений подвижных критических особых точек (свойство Пенлеве).

Для анализа решений систем (1) – (44) использован также подход, заключающийся в замене каждой из них эквивалентным уравнением третьего порядка (или второго порядка с произвольной постоянной) и сравнением его с известными уравнениями Пенлеве-типа.

III. ВЫВОДЫ

Теорема 1. Системы (2), (3), (9), (14), (16), (20), (21) не являются системами Пенлеве-типа

Доказательство. Третья компонента перечисленных выше систем имеет вид $z = Ce^{\varepsilon t}$, где C – произвольная постоянная. В силу этого системы (2), (3), (9), (14), (16), (20), (21) эквивалентны неавтономным уравнениям второго порядка [1]

$$\ddot{y} - y^2 - Ay = Ce^{\varepsilon t}, \quad (45)$$

$$\ddot{y} - y^2 - Ay = Ce^{\varepsilon t}, \quad (46)$$

$$\ddot{x} - x^2 - Ax = \varepsilon Ce^{\varepsilon t}, \quad (47)$$

$$\ddot{y} - y^2 - Ay = (\varepsilon A - A^2)Ce^{\varepsilon t}, \quad (48)$$

$$\ddot{y} - y^2 - Ay = \varepsilon Ce^{\varepsilon t}, \quad (49)$$

$$\ddot{y} - y^2 = (A + 1)Ce^{\varepsilon t}, \quad (50)$$

$$\ddot{y} - y^2 = (A + \varepsilon)Ce^{\varepsilon t} \quad (51)$$

соответственно. Сравнение уравнений (45)–(51) с уравнениями из списка [3] показывает, что они не являются уравнениями Пенлеве-типа.

Теорема 2. Системы (4), (5), (8), (10), (12), (15), (17), (23), (25)–(30), (32), (38)–(43) не являются системами Пенлеве-типа.

Проведен Пенлеве-анализ уравнения третьего порядка, отличного от уравнения

$$\ddot{u} = P(u, \dot{u}, \ddot{u}),$$

где P – полином относительно u, \dot{u}, \ddot{u} с постоянными коэффициентами, которому эквивалентна система (33).

Система (7) эквивалентна системе

$$\ddot{x} = \varepsilon \dot{x} + x^2 + Ax, \dot{z} = x. \quad (52)$$

Она не является системой Пенлеве-типа, т.к. общее решение ее первого уравнения согласно [4] содержит подвижные критические особые точки.

Выполнен Пенлеве-анализ систем (1), (6), (19), (35), (36) демонстрирующих, согласно [1], хаотическое поведение.

IV. СПИСОК ЛИТЕРАТУРЫ

- Zhang, Fu. Chaotic and nonchaotic behavior in three-dimensional quadratic system: 5 –1 dissipative cases / Fu Zhang, J. Heidel // Inter. Journal of Bif. and Chaos. – 2012. – Vol. 22, № 1. – P. 1250010.
- Грицук, Е. В. К теории нелинейных систем дифференциальных уравнений со свойством Пенлеве / Е. В. Грицук, В. И. Громак // Весці НАН Беларусі. Серыя фіз.-мат. навук. – 2010. – № 3. – С. 25–30.
- Айнс, Э. Л. Обыкновенные дифференциальные уравнения / Э. Л. Айнс // Харьков: ОНТИ, 1939. – 720с.
- Cosgrove, C. M. Chazy classes IX –XI of third-order differential equations / C. M. Cosgrove // Stud. Appl. Math. – 2001. – Vol. 104, № 3. – P. 171 –228.

AUTOMATION OF LAYOUT AND PREPARATION FOR PRINTING GRAPHIC MATERIALS FOR BOARD GAMES

Zhelenok D. A., Savenko A. G.

Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: savenko@bsuir.by

The paper presents a software tool designed to automate the layout and preparation for printing of graphic materials of board games. The software has a graphical and console interface, a well-thought-out architecture that allows you to avoid duplicating the source code of algorithms for different interfaces. The main functions of the software tool are: preparing a specification file with a wide selection of parameters (various types of printing materials, sheet formats (including roll printing), options for source files of board game materials and ways to compare them, choice of packaging strategy); generation of documents for printing in accordance with specifications. Software algorithms solve problems of two-dimensional packaging and guillotine cutting. The use of software will reduce the costs of preparing and producing board games.

INTRODUCTION

Today, board games are a common hobby that forms an entire industry. They have a wide variety both in design and theme, as well as in genres, ranging from abstract and logic games to strategy and role-playing games. There are many communities, clubs, and online resources dedicated to board games. The process of developing new games or additions to existing ones is most often iterative. This means that before creating the final version of the product, a prototype is developed, the ideas and mechanics are tested on it through the game itself, then edits are made, the ideas are tested again, etc. Each prototyping and testing cycle is a separate iteration of game development. After developing a polished version of the product, the author can contact publishers to present them with his work and, possibly, publish a game ready for sale through them. He can also try to publish it himself, working directly with printing houses. Or, as an alternative, release its materials in digital form, either free or for a fee through specialized resources. In addition, the next re-release of board games is often accompanied by changes to the rules and mechanics of the game. Changes in the rules are also often accompanied by changes in both the information contained on the game components and the components themselves. An iterative approach and the release or re-release of a board game in any form implies the layout of components for their further printing. Simplifying and increasing the quality of layout allows you to reduce the cost of publishing board games. This, in turn, reduces the cost of individual copies, which increases accessibility to reach a larger audience. Different materials are used to make different types of components. Different components require different preparation procedures for further manufacturing. The preparation of components for production differs between the development and publication stages. So, due to the iterative nature of the development process, changes are made frequently, but they are

aimed primarily at the rules and game mechanics, and not at the design of components (although this can be thought through and changed at this stage). At the publishing stage, the design must acquire its final form, which will be convenient both for production and for the end consumer. In the context of board games, the layout process involves arranging game elements so that they are easier to print and cut. At the same time, the paper format and the layout of objects depend on where the layout objects are subsequently produced.

I. PROPOSED SOLUTION

The developed software allows you to automate the process of layout and preparation of board game materials for printing. The PS supports a hybrid operating mode, which means support for both graphical and text (console) interface options. The graphical operating mode provides the user with a convenient interface for creating a specification file and its use for generating documents. The specification file is in YAML format, which in declarative form describes the parameters that are expected from the generated documents.

The specification file describes the following parameters:

- the structure of folders in which the results of the program will be saved;
- the number and name of documents that the user expects to receive;
- paper format and indents that must be observed in each document;
- source of images of game components and their expected size;
- strategy for packing objects on each page;
- source of cover images, if the component requires them;
- a method for forming cutting marks, if necessary.

Each of the points above allows you to customize the expected result. The program, reading the specification and other necessary files, can provide the following functions:

- arrangement of images on a two-dimensional limited area, taking into account indents, expected sizes of components and the packaging strategy chosen by the user;
- scaling input images, as well as forming a frame around them if it is necessary to avoid problems associated with paper displacement during printing;
- slicing input images if they do not fit into the expected area;
- placing marks for cutting;
- генерация и расстановка рубашек для тех компонентов, которым она необходима;
- the ability to distribute components both within one document and in different ones.

At the output, the program generates a structure of folders containing generated files in the Portable Document Format (PDF). The folder structure is also described in the original specification. Packing is a class of combinatorial geometry problems that involve placing figures of a given size or shape inside another given figure with the greatest economy or with some other restrictions [1]. Packaging strategy refers to the algorithm by which game components will be placed on document pages. In the developed software, packaging algorithms are divided into two main groups: optimizing the space used and simplifying post-printing processing. The first are algorithms that solve the two-dimensional packing problem (2DBP). They must be able to handle components of different sizes and their result must, if possible, minimize the space taken up on the sheets and the number of sheets of paper used. The latter must minimize the number of guillotine cuts that need to be made to cut out all the components. This problem is also known as the guillotine cutting problem. The problem solved by two-dimensional packaging algorithms is formulated as follows: given rectangles of a certain size (width and height are specified), which must be placed on larger rectangles (containers). The main goal is to find such an arrangement of all rectangles so that they do not intersect and occupy the least number of containers [2]. The two-dimensional packing problem has two options for input data: when the set of objects to be packed is known in advance (offline problem) and when the data arrives in portions (online problem). The developed software uses algorithms that solve the offline problem, since the size of all game components is known in advance. Another variation of the two-dimensional packaging problem, which is algorithmically implemented in this software, is the 2-Dimensional Strip Packing (2DSP) problem. This option differs from 2DBP in that packaging is carried out in only one container, which is limited in width, but not limited in height. Instead of minimizing the number of containers, we minimize the filling height of one [3]. The guillotine

cutting problem is a problem of combinatorial geometry, close to the cutting problem and container packing problems. The main goal is to obtain the maximum number of rectangular-sized sheets from a larger sheet by making only guillotine cuts, that is, straight cuts from edge to edge. The functions of the software tool can be conditionally divided into two categories: one allows you to create a specification that describes how to prepare the source data and convert it into a document ready for printing, the second allows you to convert the input materials into the final document. To transform a document, primitives such as point, size, and indentation are used. These in turn are combined into more complex entities: plain paper, roll paper and components. Packing objects in a two-dimensional limited space is carried out by packing strategies. To structure the model, the strategy design pattern is used. Strategy is a behavioral design pattern; it defines a family of algorithms, encapsulates each of them, and makes them interchangeable. The strategy allows you to change algorithms regardless of the clients who use them. Packing strategies return a packed document. It describes where which primitives should be located. This approach allows you not to be tied to a specific format for export, implementing it separately. The paper, component set, and packaging strategy enable the document specification to be represented. After creating a packaged document, you need to convert it to a file or display it. This again uses the strategy design pattern, but this time it describes classes for different rendering methods. Before packing the components, preparation is carried out. It involves reading and converting input images of the front and back of the components.

II. CONCLUSION

A software tool was developed to automate the layout and preparation for printing of materials for board games, solving the problem of optimal packaging in two-dimensional space and the subsequent problem of guillotine cutting. The developed software can be used both in production by board game publishers and by private users to prepare and print board game components using the PNP (print and play) method or develop their own games. The use of the developed software will reduce the costs of preparing and publishing or reprinting board games and significantly simplify the process of preparation and production.

BIBLIOGRAPHY

1. Britanica Packing Definition. [Electronic resource]. – Access mode: <https://www.britannica.com/science/packing-combinatorics>. Access date 10.10.2023.
2. Nikhil Bansal, Maxim Sviridenko. New Approximability and Inapproximability Results for 2-Dimensional Bin Packing / Maxim Sviridenko Nikhil Bansal. – 2003.
3. Хабр: Про двумерную упаковку: offline алгоритмы. [Electronic resource]. – Access mode: <https://habr.com/en/post/136225>. Access date 10.10.2023.

Секция «Автоматизированные
системы обработки информации»

Section «Automated information
processing systems»

СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ В ЗАДАЧАХ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Адамов Н. А., Шилин Л. Ю., Навроцкий А. А.

Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: adamov.nikooo@gmail.com, dekfitu@bsuir.by

В докладе рассмотрены подходы, которые позволяют увеличивать производительность сверточных нейронных сетей путем комбинирования различных эвристик.

ВВЕДЕНИЕ

Современные нейронные сети отличаются высокой производительностью и гибкостью. Высокое качество обуславливается применением различных подходов, которые позволяют значительно увеличить глубину сети, что, в свою очередь, напрямую влияет на конечный результат.

I. МЕТОДЫ ПОВЫШЕНИЯ КАЧЕСТВА ОБУЧЕНИЯ

Затухающий градиент при большой глубине нейронной сети, является проблемой т.к. не позволяет полноценно проводить обучение. Коэффициенты обучения во время обратного распространения ошибки имеют тенденцию к уменьшению, что влечет за собой незначительное изменение в весах нейронной сети и не позволяет полноценно обучать сеть. Данную проблему может решить технология, которая использовалась в сети ResNet на соревновании ImageNet по классификации объектов на 1000 классов. Данная технология называется остаточным обучением. На рисунке 1 представлен график производительности сети в зависимости от глубины. Слева – обычная сеть, справа – сеть ResNet. Обычная сеть показывает меньшую ошибку при меньшем количестве слоев в сети, однако сеть ResNet показала меньшую ошибку при большем количестве слоев, что свидетельствует об эффективности использования остаточного обучения в глубоких сетях.

Для повышения скорости сходимости градиента используются различные алгоритмы оптимизации. Алгоритм Adam позволяет найти глобальный минимум функции потерь, тем самым уменьшая ошибку модели на данных. На рисунке 2 представлены результаты тестирования алгоритмов оптимизации на основе градиентов: слева SGD, справа Adam с добавлением остаточного обучения (RL) и пакетной нормализации (BN),

а также без этих параметров. Алгоритмы SGD и Adam позволяют сети с остаточным обучением и пакетной нормализацией получать хорошие результаты. Остаточное обучение выигрывает в сравнении с пакетной нормализацией. Это связано с тем, что пакетная нормализация снижает для SNN значимость проблемы внутреннего ковариационного сдвига. Остаточное обучение без пакетной нормализации имеет высокую сходимость, однако уступает остаточному обучению с пакетной нормализацией. С другой стороны, пакетная нормализация выигрывает при использовании остаточного обучения. Без остаточного обучения пакетная нормализация имеет определенное неблагоприятное влияние на конвергенцию. При остаточном обучении пакетная нормализация может быть использована для ускорения обучения и повышения производительности. Без остаточного обучения интенсивность входных данных и сверточный признак коррелируются с соседними, а распределение входных данных слоя зависит от содержимого изображений в каждой обучающей батче. Остаточное обучение позволяет улучшить пакетную нормализацию путем уменьшения внутреннего ковариационного сдвига.

Интеграция различных подходов, например, остаточного обучения и пакетной нормализации может не только ускорить и стабилизировать процесс обучения, но и повысить эффективность модели.

СПИСОК ЛИТЕРАТУРЫ

1. Karen Simonyan, Andrew Zisserman, 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
2. Geoffrey Hinton, Nitish Srivasta, Kevin Swersky. 2012. «Lecture 6a Overview of Mini - Batch Gradient Descent) www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf. Accessed 21 Mar. 2023.

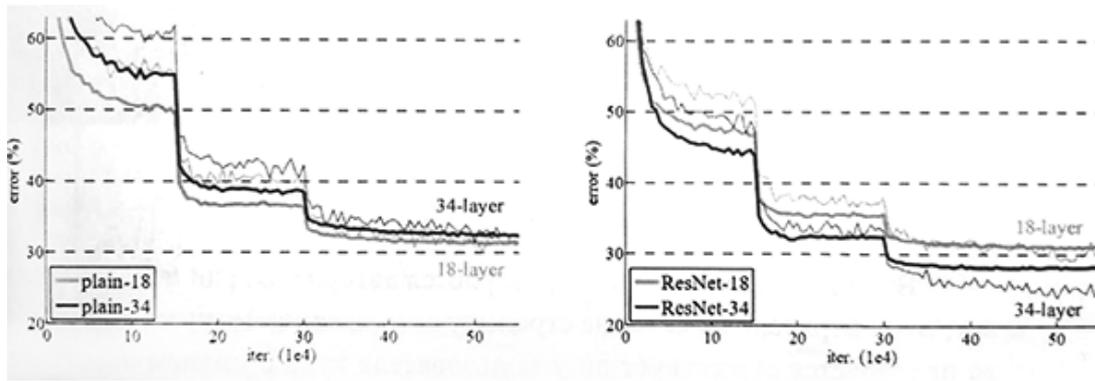


Рис. 1 – Производительность сетей

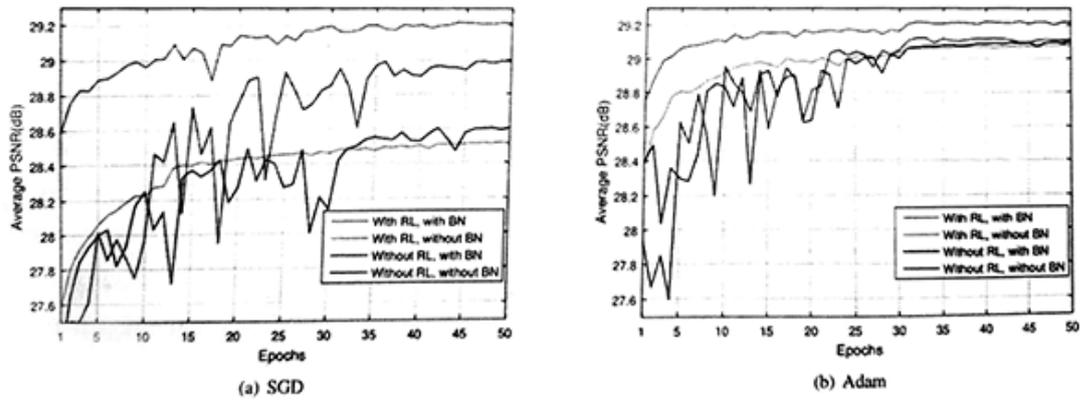


Рис. 2 – Использование алгоритмов оптимизации

АЛГОРИТМ ДЕЙКСТРЫ ПРИ ПРОЕКТИРОВАНИИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ

Акулич В. Н., Шилин Л. Ю., Хмыз Д. Д.

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

E-mail: uladzislau.akulich@gmail.com, dekfitu@bsuir.by

Разработана методика применения алгоритма Дейкстры для оптимизации автоматизированных систем. Разработана автоматизированная система с использованием алгоритма Дейкстры для нахождения кратчайшего пути и визуализации его для пользователей.

ВВЕДЕНИЕ

Применение алгоритмов теории графов при проектировании автоматизированных систем позволяет реализовывать новые функции и возможности систем. Использование систем, использующих алгоритмы теории графов, является значительным шагом вперёд для многих организаций и пользователей. Алгоритм Дейкстры является важным инструментом для оптимизации и планирования в различных областях, включая разработку автоматизированных систем.

ПРИМЕНЕНИЕ АЛГОРИТМА ДЕЙКСТРЫ

Алгоритм Дейкстры – это метод, который находит кратчайший путь от одной вершины графа к другой. Граф – структура из точек-вершин, соединённых ребрами. Алгоритм Дейкстры работает для графов, у которых нет ребер с отрицательным весом, то есть таких, при прохождении через которые длина пути уменьшается.

В отличие от похожих методов, алгоритм Дейкстры ищет оптимальный маршрут от одной заданной вершины ко всем остальным, при этом высчитывая длину пути – суммарный вес ребер, по которым проходит при этом маршруте.

Применение алгоритма Дейкстры при проектировании автоматизированных систем является эффективным, так как алгоритм может быть использован для решения следующих задач:

- автоматическое построение маршрута на карте;
- моделирование движения робота;
- маршрутизация движения данных в компьютерной сети;
- обработка транспортных потоков.

Рассмотрим алгоритм Дейкстры для построения кратчайшего пути и визуализации его на карте.

Задача о кратчайшем пути – задача поиска самого короткого пути между двумя вершинами на графе, в котором минимизируется сумма весов ребер, составляющих путь. В качестве вершин на карте выступают перекрестки, здания и другие объекты, а дороги, которые лежат между ними, являются ребрами. Если сумма длин дорог меж-

ду перекрёстками минимальна, тогда найденный путь самый короткий.

Задан ориентированный граф $G(V,E)$, каждой дуге (u,v) ставится в соответствие число $L(u,v)$ – длина дуги (расстояния, стоимости). В общем случае возможно $L > 0$, $L < 0$, $L = 0$.

Под длиной пути понимаем сумму длин дуг, составляющих путь.

Главная задача: найти длины кратчайших путей и сами пути от фиксированной вершины s до всех остальных вершин графа v_i .

Алгоритм решает задачу в случае $L \geq 0$. Алгоритм основан на приписывании вершинам v_i временных меток $d(v_i)$. Метка вершины даёт верхнюю границу длины пути от s к этой вершине. Величины меток постепенно уменьшаются, и на каждом шаге итерации одна из временных меток становится постоянной. Это означает, что метка даёт точную длину кратчайшего пути от s к рассматриваемой вершине.

Согласно алгоритму Дейкстры процесс построения дерева начинается с заданной вершины. Положить $d(s)=0$, считать метку постоянной. $d(v_i) = \infty, i = 1..n$, считать метки временными. $p=s$.

В дальнейшем на каждом шаге к дереву присоединяется одно новое ребро (и одна вершина). Это ребро выбирается из подходящих ребер, причем подходящим считается ребро, соединяющее вершину дерева с вершиной, ему не принадлежащей. Среди подходящих ребер выбирается ребро наименьшего веса.

Повторяется n раз, пока не будут упорядочены все вершины. Временная метка $d(v_i)$ всякой неупорядоченной вершины v_i , в которую входит дуга, выходящая из p , считается по формуле:

$$d(v_i) = \min(d(v_i), d(p) + l(pv_i))$$

Выбирается вершина с $\min(d(v_i))$ Если их несколько, то любую. Пусть это w . Метку $d(w)$ считать постоянной. $p=w$.

Доказательство корректности самого алгоритма основывается на следующем утверждении. После того, как какая-либо вершина v становится помеченной, текущее расстояние до неё d_v уже

является кратчайшим, и, соответственно, больше меняться не будет.

Реализация алгоритма Дейкстры в автоматизированной системе, в которой автоматически вычисляется и отображается кратчайший путь на карте, состоит из следующих этапов:

1. Инициализация – все вершины графа обозначаются как недостижимые, а расстояние от начальной вершины до всех остальных вершин устанавливается как бесконечность, кроме начальной вершины, для которой устанавливается расстояние 0.
2. Создание пустого множества, которое будет содержать посещенные вершины.
3. Выбор начальной вершины, из которой будет начинаться процесс поиска кратчайших путей.
4. Обновление расстояний обновление расстояний до соседних вершин - для каждой соседней вершины, которая еще не была посещена, вычисляется новое расстояние, если оно оказывается меньше, чем текущее расстояние до этой вершины, расстояние обновляется.
5. Пометка текущей вершины как посещенной.
6. Поиск следующей вершины для посещения - среди всех непосещенных точек на основе их расстояний от начальной вершины выбирается вершина с наименьшим расстоянием.
7. Повторение этапов с выбором вершин и расчетом расстояния, пока все вершины не будут посещены или пока не будет найден кратчайший путь до заданной конечной вершины.
8. Формирование кратчайшего пути.

Рассмотрим пример задачи для нахождения кратчайшего пути в ориентированном графе. Ориентированный граф представлен на рисунке 1.

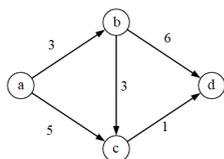


Рис. 1 – Ориентированный граф

В условиях решения задач автоматизированных систем с использованием онлайн-карт рёбра ab , bc , ac , bd , db будут выступать в качестве до-

рог, а вершины a , b , c , d – в качестве зданий или других объектов.

Первый шаг алгоритма определит, что кратчайший путь от вершины a будет проходить в вершину b и зафиксирует кратчайший путь.

$$l_{ab} = 3;$$

$$l_{ac} = 5;$$

Второй шаг рассмотрит все возможные варианты $A[v] + l_{vw}$. Получим значения:

$$l_{ab} + l_{bc} = 3 + 3 = 6;$$

$$l_{ab} + l_{bd} = 3 + 6 = 9;$$

Оптимальный шаг двигаться по ребру ac , так как $l_{ac} < l_{ab} + l_{bc}$ и $l_{ac} < l_{ab} + l_{bd}$.

Добавляется длина кратчайшего пути от вершины a до вершины c .

На третьем шаге находим $l_{ac} + l_{cd}$.

$$l_{ac} + l_{cd} = 3 + 1 = 4;$$

Таким образом, благодаря алгоритму Дейкстры длина всех путей будет просчитана и будет выбран кратчайший путь до любой из вершин.

Одним из главных преимуществ алгоритма Дейкстры является его значительно низкая сложность, которая является почти линейной, поэтому его внедрение оптимизирует систему.

I. ЗАКЛЮЧЕНИЕ

Алгоритма Дейкстры в автоматизированных системах применяется для решения разных задач, связанных с построением кратчайшего пути, что позволяет определить оптимальные маршруты с учетом самых различных факторов. Использование алгоритма Дейкстры является эффективным в транспортных и геоинформационных системах, сетях связи, играх с искусственным интеллектом, а также в других сферах применения автоматизированных систем.

СПИСОК ЛИТЕРАТУРЫ

1. Рафгарден, Т. Совершенный алгоритм. Графовые алгоритмы и структуры данных. / Т. Рафгарден. – СПб.: Питер, 2019. – 256 с.
2. Карпов, Д. В. Связность графов : / Д. В. Карпов. – СПб.: Санкт-Петербургское отделение Мат. инст. им. В. А. Стеклова РАН, 2018. – 183 с.

СНИЖЕНИЕ РИСКОВ УТЕЧКИ ПЕРСОНАЛЬНЫХ ДАННЫХ ПУТЕМ ИХ ОБЕЗЛИЧИВАНИЯ И АНОНИМИЗАЦИИ В СИСТЕМАХ ЭЛЕКТРОННОГО ОБУЧЕНИЯ

Барановский О. К.

Открытое акционерное общество «АГАТ – системы управления» –
управляющая компания холдинга «Геоинформационные системы управления»

Минск, Республика Беларусь
E-mail: baranovskiy-ok@agat.by

Рассмотрен вопрос повышения конфиденциальности пользователей систем электронного обучения в условиях регистрации их цифрового следа. Предложен подход по совместному применению методов обезличивания и анонимизации персональных данных при сохранении возможности их безопасного использования, в том числе для целей повышения качества образовательного процесса.

ВВЕДЕНИЕ

Существенный вклад в развитие систем электронного обучения (далее – СЭО) на основе результатов анализа больших данных, представляющих собой цифровой след обучающихся, определяется применением аналитических методов выявления скрытых (изначально неформализованных) закономерностей в сгенерированных массивах образовательных данных. Соответственно, СЭО являются инструментом обработки широкой номенклатуры персональных данных (далее – ПДн) обучающихся. И если перечень ПДн, получаемых от субъекта для целей оказания образовательных услуг, однозначно закрепляется документально до момента их сбора, то согласование с субъектами ПДн полного перечня генерируемых в СЭО ПДн, в том числе с использованием интеллектуального машинного анализа цифрового следа субъектов ПДн, является редкой практикой у операторов ПДн. По мнению экспертов, к недопустимым событиям информационной безопасности в учреждениях образования, эксплуатирующих СЭО, является кража ПДн субъектов образовательного процесса [1]. По сравнению с отказами в обслуживании СЭО, утечка массивов данных с ПДн наносит более высокий ущерб репутации учреждений образования.

I. ЦЕЛЬ

Приоритетной задачей при обработке ПДн пользователей СЭО является защита их от утечки с последующей потерей конфиденциальности. В этой связи наряду с выполнением обязательных требований законодательства по защите ПДн от несанкционированного доступа (далее – НСД), оператор ПДн (владелец СЭО) реализует иные меры на соответствие требованиям, которые декларируются законодательством как рекомендуемые [2]. К техническим мерам защиты информации относят реализацию методов обезличивания ПДн в прикладном программном обеспечении. Целью работы является установление подхода по обезличиванию и анонимизации ПДн в СЭО в

целях снижения рисков их утечки в результате осуществления НСД к данным в СЭО с их последующей выгрузкой вследствие отсутствия или обхода реализованных мер защиты информации.

II. ПОДХОД К ОБЕЗЛИЧИВАНИЮ И АНОНИМИЗАЦИИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Будем учитывать, что законодательство Республики Беларусь определяет обезличивание ПДн как действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность ПДн конкретному субъекту ПДн. Это означает, что для обеспечения функционирования систем обработки ПДн применяются средства восстановления (деобезличивания) ПДн. При этом исследователи в сфере информационной безопасности неоднократно доказывали возможность повторной неправомерной персонализации обезличенных ПДн вследствие некорректного применения методов обезличивания.

Законодательством регламентированы методы обезличивания, подлежащие применению в системах защиты информации информационных систем: введение идентификаторов, изменение состава, декомпозиция, перестановка, зашифрование [2]. Методы обезличивания применяются для защиты ПДн при их хранении. Очевидно, что для заданной модели угроз (модели атак) ПДн выбор методов обезличивания и их комбинация будет определяться не менее чем тремя факторами: категориями ПДн (определяют уровень ущерба для субъектов ПДн); информационными технологиями обработки, хранения, передачи и удаления ПДн; показателями производительности основных операций над ПДн, реализующих функции назначения СЭО.

Например, при использовании СУБД выбор и комбинация методов обезличивания будет определяться приоритетами в номенклатуре операций над записями: быстрый поиск, каскадное изменение или удаление, иные сочетания групповых операций, формирование обезличенных или персонализированных выборок. Использование ма-

шинной обработки ПДн всегда предпочтительней ручной обработки администратором или пользователем с высокими привилегиями в СЭО.

Выбор и вычислительная сложность реализации методов обезличивания должны определяться путем обработки рисков, связанных с утечкой ПДн. Экспертами принято, что потеря конфиденциальности ПДн, содержащими сведения об академической успеваемости, оценках и результатах аттестации оценивается как средний уровень ущерба, поскольку ожидается, что в этом случае субъекты ПДн столкнутся со значительными неудобствами [3]. Вместе с тем, данные, накопленные в результате взаимодействия обучающегося с СЭО, (цифровой след) могут содержать сведения, если не явно выявляющие, то позволяющие извлекать информацию о когнитивных, творческих, коммуникативных способностях субъектов, их политических взглядах, религиозных и других убеждениях, а также биометрические поведенческие характеристики. Соответственно, при оценке риска необходимо принимать высокий уровень ущерба.

При формировании перечня ПДн, следует учитывать следующее. Во-первых, СЭО, которые, как правило, строятся на основе проприетарного или свободно распространяемого ПО, ведут запись событий о функционировании ПО и действиях пользователей в журналы. Платформа Moodle журналирует связанные с пользователями события в таблицу базы данных: дата и время события, IP-адрес компьютера пользователя, фамилия, имя и отчество пользователя, действие, выполненное пользователем, название и компонент события, дополнительная информация. Во-вторых, ПДн могут генерироваться при анализе больших данных, выгруженных из СЭО и являющихся цифровым следом при взаимодействии обучающихся со средствами аудио-, видеоконференцсвязи, средствами совместной работы с учебным и справочным материалом в графической и текстовой формах.

Между тем, согласно законодательству [4], содержание и объем обрабатываемых ПДн должны соответствовать заявленным целям их обработки. Обрабатываемые ПДн не должны быть избыточными по отношению к заявленным целям их обработки. Соответственно, если цели обработки – оказание образовательных услуг, то повышение их качества или развитие СЭО – это иная цель и должно быть получено согласие обучающегося (он имеет право отказаться) с предоставлением полного списка собираемых ПДн.

Вместе с тем, не требуется получать согласие владельца ПДн, если ПДн обрабатываются на основании законодательства. К подлежа-

щим регистрации содержащих ПДн событиям, связанным с защитой информации и не требующим согласия на их сбор, можно отнести идентификацию и аутентификацию пользователей, нарушение прав доступа пользователей, выявленные нарушения информационной безопасности и др. [2], также для прикладного программного обеспечения: аутентификация (вход и (или) выход) пользователей, успешные и неуспешные попытки аутентификации; неудавшиеся или отмененные действия пользователей; действия пользователей (доступ к объекту (данным), изменения объекта (данных), удаление объекта (данных)).

Для целей повышения качества образовательных услуг предпочтительней использовать методы анонимизации ПДн, что предотвратит несанкционированный поиск персонализированной информации в больших данных (запись в журнале считается анонимной, если ее данные, по отдельности или в сочетании с другими данными, не могут быть связаны с конкретным субъектом ПДн).

В этой связи операторы ПДн (владельцы СЭО) принимают решение о номенклатуре собираемых ПДн для целей оказания образовательных услуг в соответствии с законодательством, а также о номенклатуре ПДн, немедленно анонимизируемых после их получения с дальнейшим их использованием для целей повышения качества образовательных услуг и развития СЭО.

ЗАКЛЮЧЕНИЕ

Система образования Республики Беларусь находится на этапе, когда проблема обеспечения конфиденциальности обучающихся при использовании СЭО уже замечена, но для ее решения принято недостаточно мер. Учреждения образования с разным темпом вырабатывают практики снижения рисков утечки ПДн при автоматизированной обработке образовательных данных.

1. Positive Research / 2023. Сборник исследований по практической безопасности [Электронный ресурс] – Режим доступа: <https://ptresearch.media/>. – Дата доступа: 26.09.2023.
2. О мерах по реализации Указа Президента Респ. Беларусь от 9 декабря 2019 г. № 449 [Электронный ресурс] : приказ Оперативно-аналитического центра при Президенте Респ. Беларусь, 22 февраля 2020 г., № 66 // ЭТАЛОН. Законодательство Респ. Беларусь / Нац. центр правовой информ. Респ. Беларусь. – Минск, 2023.
3. Alier, M. Privacy and E-Learning: A Pending Task / M. Alier, M. J. Casan Guerrero, D. Amo, C. Severance, D. Fonseca // Sustainability. – 2021. – Vol. 13. – № 16. – P. 9206.
4. О защите персональных данных [Электронный ресурс] : Закон Республики Беларусь, 7 мая 2021 г., № 99-З // ЭТАЛОН. Законодательство Респ. Беларусь / Нац. центр правовой информ. Респ. Беларусь. – Минск, 2023.

АНАЛИЗ УЯЗВИМОСТЕЙ ПОТОКОВЫХ КРИПТОСИСТЕМ НА БАЗЕ М-ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Болтак С. В., Русакович Е. С.

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

Исследована криптостойкость потоковых криптосистем на базе М-последовательностей. Показана уязвимость данных систем. Предложен алгоритм, основанный на оптимизации метода Касиски. Разработано программное средство для взлома потоковых шифров.

ВВЕДЕНИЕ

Характерной особенностью потоковых криптосистем является использование криптографического ключа большой длины, равного длине шифруемого сообщения, для чего применяются генераторы псевдослучайных последовательностей [1].

М-последовательность – псевдослучайная двоичная последовательность максимальной длины, сгенерированная регистром сдвига с линейной обратной связью (Linear Feedback Shift Register – LFSR).

Генераторы М-последовательностей являются одним из наиболее часто используемых типов генераторов, так как обладают хорошими статистическими свойствами, которые практически не отличаются от свойств случайных последовательностей, и воспроизводимостью [2].

I. АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

Для потоковых систем шифрования наиболее известными являются следующие методы криптоанализа: метод полного перебора ключей, метод Берлекэмпа-Мэсси, побочные атаки, корреляционные методы. Полный перебор требует больших затрат по времени, а также вычислительных мощностей и при большой длине ключа практически невозможен, так как основан на переборе всех возможных ключей и одновременного дешифрования. Алгоритм Берлекэмпа-Мэсси строит эквивалентный линейный рекуррентный регистр (ЛРР) с помощью которого генерируется исходная ключевая последовательность. Однако для построения ЛРР необходимо знать часть исходной информации и соответствующую ей криптограмму. Побочные атаки возможны при повторении шифруемых битов в ключевой последовательности, когда они повторяются и шифруют разную информацию, таким образом возникает уязвимость и шифр может быть вскрыт. Корреляционные атаки – наиболее распространённые атаки, что обусловлено построением поточных шифров. Они направлены на восстановление начального состояния регистров и используют корреляцию между их входными состояниями и выходной (ключевой) последовательностью схемы [3].

LFSR в качестве генераторов М-последовательностей имеют уязвимость, которая обусловлена линейным характером генерируемой последовательности. Если злоумышленнику удалось завладеть парой «исходный текст – шифротекст» длиной $2m$ бит, где m – длина регистра, то он может восстановить характеристический многочлен и, как результат, дешифровать сообщение [1].

Использование метода Касиски предполагает наличие периодического ключа. Суть метода заключается в том, что одинаковые биты шифруемой информации дают одинаковые фрагменты (1-граммы) криптограмм, если были зашифрованы одной и той же битовой ключевой последовательностью. Однако данный метод не исключает случайные 1-граммы, что часто затрудняет определение периода ключа и позволяет определить его лишь с той или иной вероятностью [1].

При большом объёме шифруемой информации ключевая битовая последовательность будет иметь повторения, что делает возможным использование метода Касиски.

II. РЕАЛИЗАЦИЯ

Предлагаемый алгоритм состоит из двух частей – определения расстояний между повторяющимися 1-граммами и анализа расстояний для нахождения длины М-последовательности.

Для определения расстояний алгоритм выделяет первую 1-грамму, ищет ее повторения до конца файла и заносит их в специальную структуру данных. Для увеличения быстродействия длина 1-граммы берётся 5 байт, а поиск происходит многопоточно. В результате поиска будет сформирована структура, хранящая повторяющиеся 1-граммы и соответствующие им расстояния, через которые данная 1-грамма повторилась. Основная часть алгоритма – это анализ найденных расстояний и нахождение длины ключа. Классический тест Касиски предполагает нахождение наибольшего общего делителя найденных расстояний (НОД), это и есть предполагаемая длина периода ключа. Схема алгоритма функции анализа найденных расстояний между повторяющимися 1-граммами изображена на рисунке 1.

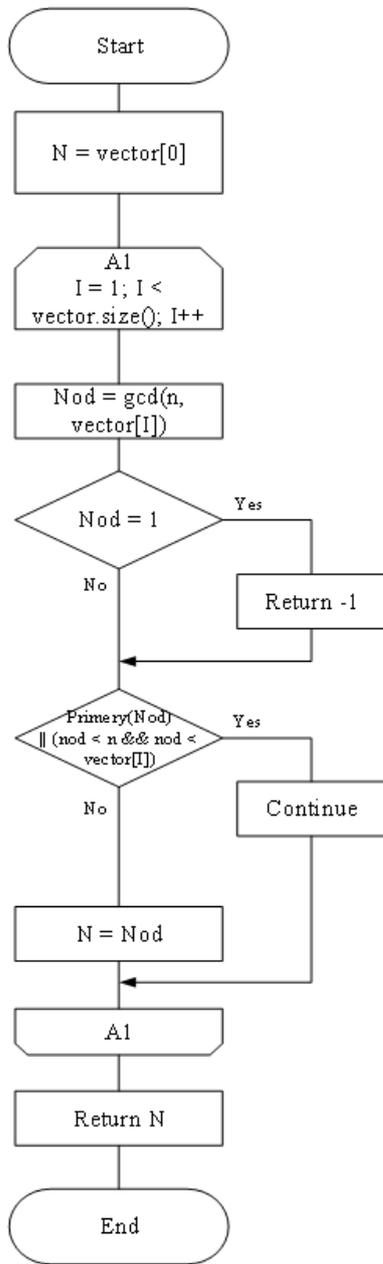


Рис. 1 – Алгоритм анализа расстояний

Для более точного нахождения длины ключа в существующий метод были внесены некоторые ограничения. Задача анализа расстояний состоит в том, чтобы исключить те расстояния, которые могут привести к нарушению корректности нахождения длины ключа, так как если хотя бы одно из расстояний случайно, оно изменит результат вычисления длины ключа для всех остальных расстояний. Например, если это случайное число простое, то нахождение НОД с этим числом приведёт к получению либо 1, либо самого числа, где 1 — всегда неверно, то есть вероятность правильного ответа низкая. Следовательно, необходимо на уровне анализа одной l-граммы получить непростое число с максимальным количеством делителей (чтобы повысить вероятность

нахождения истинной длины ключа в делителях на дальнейшем этапе).

Шаги алгоритма:

1. взять первое расстояние;
2. взять следующее расстояние;
3. найти НОД двух расстояний;
4. если НОД равен 1, записать в результирующий массив -1;
5. если НОД не равен 1, то шаг 6;
6. если полученное значение не является простым числом, а также не равно ни одному из чисел, участвующих в вычислении НОД – текущему расстоянию присвоить найденный НОД;
7. повторять шаги 2-6 пока в массиве расстояний не останется одно число.

Описанный выше алгоритм применяется для всех найденных ранее l-грамм. В результате каждому массиву расстояний будет поставлено в соответствие одно число. Далее из массива удаляются все -1 и находится НОД оставшихся чисел. Найденное число и будет являться длиной искомого ключа.

III. ВЫВОДЫ

Тестирование работы предложенного алгоритма проводилось для шифротекстов, полученных с помощью M-последовательностей на базе примитивных многочленов степени два, четырнадцать, восемнадцать и двадцать три. Время, понадобившееся для взлома, отображено в таблице 1.

Таблица 1 – Время взлома

Степень полинома	Время взлома, мс
2	2
4	122221
18	27843
23	3022844

Вероятность взлома для предлагаемого метода зависит от наличия повторений в шифрующей битовой последовательности, а также ограничена вычислительной мощностью используемого для вычислений компьютера.

Проведенные исследования могут быть использованы разработчиками потоковых крипто-систем.

СПИСОК ЛИТЕРАТУРЫ

1. Ярмолик, В. Н. Элементы теории информации: Практикум / В. Н. Ярмолик, А. П. Занкович, С. С. Портайко. – Минск: БГУИР, 2007. – 7–11 с.
2. Golomb, S. W. Shift register sequences. –San Francisco: Holden-Day, 1967. – 24 p.
3. Кутузов, А. В. Методы криптоанализа блочных и поточных систем шифрования / А. В. Кутузов, Н. В. Старченков, Д. А. Кудряшов // Современная техника и технологии. 2016. №9. – [Электронный ресурс]. Режим доступа: <https://technology.snauka.ru/2016/09/10571>. – Дата доступа: 12.08.2023.

МЕТАПРОГРАММИРОВАНИЕ В C# С ИСПОЛЬЗОВАНИЕМ ДЕРЕВЬЕВ ВЫРАЖЕНИЙ

Войтович А. С., Ломако А. В.

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

E-mail: voitovich@gmail.com, lavlot@bsuir.by

Анализируются деревья выражений на платформе .NET. Рассматриваются методы построения деревьев выражений, компиляция кода на этапе выполнения, использование деревьев выражений для повышения производительности.

ВВЕДЕНИЕ

Деревья выражений – малоизвестная, но очень интересная функция .NET. Большинство людей, знают их как средство получения структур объектно-реляционного отображения, но, несмотря на то, что это наиболее распространенный вариант использования, он не единственный. Деревья выражений позволяют реализовать самый разнообразный функционал, включая генерацию кода, транпиляцию, метапрограммирование и многое другое.

Целью данной статьи является обзор того, что такое деревья выражений и как с ними работать, а также тестирование сценария использования выражений для оптимизации работы с информацией о типах и пространством имен System.Reflection.

I. ПОСТРОЕНИЕ ДЕРЕВЬЕВ ВЫРАЖЕНИЙ

Когда дело доходит до языков программирования, выражение описывает некоторую операцию над данными, которая дает определенный результат. Это одна из основополагающих конструкций любого языка.

Хотя для людей, это очевидно, для программной интерпретации этого представления нам нужна специальная структура данных. Эту структуру данных называют деревом выражений.

В C# деревья выражений можно создавать двумя способами: мы можем создавать их напрямую через API, а затем скомпилировать их в инструкции времени выполнения, или мы можем дизассемблировать их из предоставленных лямбда-выражений.

Фреймворк предлагает нам API для построения деревьев выражений через класс Expression, расположенный в пространстве имен System.Linq.Expressions. Он предоставляет различные фабричные методы, которые можно использовать для создания выражений разных типов. Рассмотрим создание выражения на рисунке 1.

```
public Func<string, string?> ConstructGreetingFunction()
{
    var personNameParameter =
        Expression.Parameter(typeof(string), "personName");

    // Condition
    var isNullOrWhiteSpaceMethod = typeof(string)
        .GetMethod(nameof(string.IsNullOrEmpty));

    var condition = Expression.Not(
        Expression.Call(isNullOrWhiteSpaceMethod, personNameParameter));

    // True clause
    var concatMethod = typeof(string)
        .GetMethod(nameof(string.Concat), new[] { typeof(string), typeof(string) });

    var trueClause = Expression.Call(
        concatMethod,
        Expression.Constant("Greetings, "),
        personNameParameter
    );

    // False clause
    var falseClause =
        Expression.Constant(null, typeof(string));

    var conditional =
        Expression.Condition(condition, trueClause, falseClause);

    var lambda =
        Expression.Lambda<Func<string, string?>>(conditional, personNameParameter);

    return lambda.Compile();
}
```

Рис. 1 – Пример создания выражения

Сначала мы вызываем Expression.Parameter() для создания выражения параметра. Мы сможем использовать его для получения значения, переданного определенному параметру.

После этого мы извлекаем тип для получения ссылки на метод string.IsNullOrEmpty(). Мы используем Expression.Call() для создания выражения вызова метода, которое представляет собой вызов string.IsNullOrEmpty() с параметром, разрешенным выражением, которое мы создали ранее. Чтобы выполнить логическую операцию «нет» над результатом, мы вызываем Expression.Not(), чтобы обернуть вызов метода.

Чтобы составить положительное предложение, мы создаем операцию «добавить» с помощью Expression.Add(). В качестве операндов мы предоставляем постоянное выражение для строки «Greetings» и выражение параметра, полученное ранее.

Затем для отрицательного случая мы используем Expression.Constant() для создания нулевого константного выражения. Чтобы гарантировать правильность ввода значения null, мы явно указываем тип в качестве второго параметра.

Наконец, мы объединяем все вышеперечисленные части вместе, чтобы создать наш тер-

нарный условный оператор. Если вы потратите немного времени на то, чтобы проследить, что входит в `Expression.Condition()`, вы поймете, что мы по существу воспроизвели древовидную диаграмму, которую видели ранее.

Однако само по себе это выражение бесполезно. Поскольку мы создали его сами, нас не особо интересует его структура – вместо этого мы должны получить выражение. Для этого нам нужно создать точку входа, обернув все в лямбда-выражение. Чтобы превратить его в настоящую лямбду, мы можем вызвать метод `Compile()`, который создаст делегат, который мы можем вызвать.

II. ПОВЫШЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ С ИСПОЛЬЗОВАНИЕМ ВЫРАЖЕНИЙ

Когда дело доходит до скомпилированных выражений, одним из наиболее распространенных сценариев использования является код с использованием `reflection`. Как известно, отражение может быть довольно медленным из-за позднего связывания, однако, компилируя код во время выполнения, мы можем добиться большей производительности.

Рассмотрим класс, на рисунке 2, у которого есть закрытый метод, который необходимо вызвать извне.

```
public class Command
{
    // Ссылка 0
    private int Execute() => 42;
}
```

Рис. 2 – Класс с закрытым методом

С помощью `reflection` это сделать довольно просто, рисунок 3:

```
public static int CallExecute(Command command) =>
    (int)typeof(Command)
        .GetMethod("Execute", BindingFlags.NonPublic | BindingFlags.Instance)
        .Invoke(command, null);
```

Рис. 3 – Вызов частного метода с использованием `reflection`

Конечно, вызов такого метода может вызвать серьезные проблемы с производительностью, если мы поместим его в цикл. Прежде чем перейти к выражениям, мы можем сначала оптимизировать приведенный выше код, отделив часть, которая получает `MethodInfo`, от части, которая вызывает метод, поскольку если этот метод вызывается более одного раза, нам не нужно каждый раз использовать `GetMethod()`, рисунок 4:

```
public static class ReflectionCached
{
    // Ссылка 1
    private static MethodInfo ExecuteMethod { get; } = typeof(Command)
        .GetMethod("Execute", BindingFlags.NonPublic | BindingFlags.Instance);

    // Ссылка 1
    public static int CallExecute(Command command) =>
        (int)ExecuteMethod.Invoke(command, null);
}
```

Рис. 4 – Вызов частного метода с использованием `reflection` и кешированием

Таким образом мы получили практически максимальную производительность, насколько это возможно с использованием `reflection`. Теперь сделаем то же самое, используя скомпилированное выражение, рисунок 5:

```
public static class ExpressionTrees
{
    // Ссылка 1
    private static MethodInfo ExecuteMethod { get; } = typeof(Command)
        .GetMethod("Execute", BindingFlags.NonPublic | BindingFlags.Instance);

    // Ссылка 2
    private static Func<Command, int> Impl { get; }

    // Ссылка 0
    static ExpressionTrees()
    {
        var instance = Expression.Parameter(typeof(Command));
        var call = Expression.Call(instance, ExecuteMethod);
        Impl = Expression.Lambda<Func<Command, int>>(call, instance)
            .Compile();
    }

    // Ссылка 1
    public static int CallExecute(Command command) => Impl(command);
}
```

Рис. 5 – Вызов частного метода с использованием `Expression.Compile`

Во всех этих подходах используются статические конструкторы для инициализации свойств ленивым и потокобезопасным способом. Это гарантирует, что вся тяжелая работа произойдет только один раз – при первом обращении к членам этих классов.

Теперь сравним эти методы друг с другом и оценим их производительность с помощью `Benchmark.NET`, таблица 1:

Таблица 1 – Результаты тестирования производительности

Метод	Среднее время выполнения, нс	Потребление памяти, Б
Reflection	113.225	48
Reflection (cached)	86.783	48
Expressions	2.716	24

ЗАКЛЮЧЕНИЕ

Использование деревьев выражений позволяет значительно улучшить гибкость системы и осуществлять кодогенерацию на этапе выполнения программы. При этом, из проведенного исследования видно, что скомпилированные выражения позволяют достичь производительности, сравнимой с нативной компиляцией, при этом получив все преимущества генерации кода и снизить накладные расходы на использование `reflection`.

1. Don Syme. 2006. Leveraging .NET meta-programming components from F# integrated queries and interoperable heterogeneous execution – Proceedings of the 2006 workshop on ML, 2006.
2. Hazzard K. Metaprogramming in .NET / Hazzard K., Bock J. – New York : Manning, 2019. – 360 p.

СКОРИНГОВАЯ МОДЕЛЬ ОЦЕНКИ КРЕДИТНОГО РИСКА

Высоцкий М. В., Жилияк Н. А., Шабанович Р. А.
Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: m.vysotskij@bsuir.by, gznadya@gmail.com, r.shabanovich@bsuir.by

В данной работе рассмотрены методы, принципы, преимущества, недостатки, оценка эффективности скоринговой модели

ВВЕДЕНИЕ

Банковский сектор является одним из важнейших элементов экономики страны, обеспечивающий движение финансовых ресурсов и, соответственно, её дальнейшее развитие. Кредитный риск является одной из основных проблем, с которыми сталкиваются банки и другие финансовые учреждения. Он возникает, когда заемщик не может выплатить свой долг в срок, что может привести к финансовым потерям для кредитора. В связи с этим, оценка кредитного риска стала важной частью банковского бизнеса. Скоринговые модели играют важную роль в оценке кредитного риска. Они используются для прогнозирования вероятности невыплаты долга заемщиком на основе различных характеристик заемщика и кредита.

I. МЕТОДЫ СКОРИНГА

Методы скоринга включают различные подходы, которые могут быть классифицированы в следующие категории:

1. Статистические методы — методы, такие как многомерный дискриминантный анализ и *probit*-анализ, использующие статистические модели для оценки кредитного риска.
2. Методы управления наукой — методы, представленные математическим программированием, использующие оптимизационные модели для оценки кредитного риска.
3. Методы *data mining* — методы, такие как деревья решений и нейронные сети, использующие алгоритмы машинного обучения для оценки кредитного риска.
4. Искусственный интеллект и машинное обучение — методы, использующие сложные алгоритмы, такие как случайные леса, градиентный бустинг и глубокие нейронные сети, для оценки кредитного риска.
5. FICO Scoring Model — алгоритм, который производит то, что считается наиболее надежными кредитными рейтингами. Около 90% кредиторов используют модель FICO для оценки кандидатов на кредит.

Важно отметить, что выбор конкретного метода скоринга зависит от многих факторов, вклю-

чая доступные данные, бизнес-цели и специфические требования к регулированию.

II. ПРИНЦИПЫ ПОСТРОЕНИЯ СКОРИНГОВОЙ МОДЕЛИ

Основные принципы построения скоринговой модели включают следующие шаги:

1. Определение ключевых бизнес-целей: необходимо определить, как скоринговая модель может помочь достичь бизнес-целей организации [1].
2. Оценка доступных данных: оцените, какие данные у вас уже есть и какие данные могут быть необходимы для построения модели [1].
3. Выбор метода моделирования: методы, используемые для создания и проверки моделей кредитного скоринга, включающие логистическую и линейную регрессию, машинное обучение и предсказательную аналитику [2].
4. Разработка модели: на основе выбранного метода разрабатывается модель, которая использует собранные данные для прогнозирования кредитного риска [2].
5. Оценка модели: модель оценивается с использованием различных метрик, таких как точность, полнота, AUC-ROC [2].
6. Применение AI и ML: искусственный интеллект и машинное обучение могут помочь вам создать модель кредитного скоринга, которая соответствует вашим бизнес-целям [1].

III. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ СКОРИНГОВЫХ МОДЕЛЕЙ

Преимущества:

1. Эффективность: позволяют быстро и эффективно оценивать кредитоспособность заемщика.
2. Объективность: минимизируют человеческие ошибки и предубеждения, обеспечивая более объективную оценку.
3. Снижение риска: помогают банкам управлять кредитным риском и снижать уровень просроченной задолженности.

4. Увеличение прибыли: способствуют банкам увеличить прибыль за счет более точного определения ставок и лимитов.
5. Улучшение отношений с клиентами: улучшают отношения с клиентами за счет более справедливого и прозрачного процесса оценки.

Недостатки:

1. Недостаточная точность: скоринговые модели могут не всегда точно предсказывать поведение заемщика.
2. Дискриминация: могут быть предвзятыми и дискриминировать определенные группы заемщиков.
3. Сложность интерпретации: некоторые сложные модели, основанные на машинном обучении, могут быть трудными для понимания и объяснения.
4. Зависимость от качества данных: качество и актуальность входных данных может существенно влиять на точность модели.
5. Необходимость постоянного обновления: модели требуют регулярного обновления и адаптации к изменяющимся экономическим условиям.

IV. ОЦЕНКА ЭФФЕКТИВНОСТИ СКОРИНГОВОЙ МОДЕЛИ

Оценка эффективности скоринговой модели включает в себя использование различных методов:

1. Линейная регрессия — это статистический метод, используемый для оценки взаимосвязи между двумя или более переменными.
2. Логистическая регрессия — это метод анализа, который позволяет прогнозировать бинарный исход, на основе одной или нескольких предикторных переменных [1].
3. Предсказательная аналитика — это область статистики, которая сосредоточена на использовании статистических моделей и анализа, для прогнозирования будущих результатов [1].
4. Дискретизация — это процесс преобразования непрерывных переменных или функций в дискретную форму [1].
5. Кумулятивный профиль точности (CAP) — это графический метод оценки качества модели прогнозирования [1].

6. Рабочая характеристика приемника (ROC-кривая) — это графический метод оценки качества бинарного классификатора [1].

V. ЗАКЛЮЧЕНИЕ

В данной работе были рассмотрены скоринговые модели оценки кредитного риска. Разобраны методы, принципы построения, преимущества, недостатки и оценку эффективности скоринга. Скоринговая модель позволяет быстро, объективно и эффективно оценивать кредитоспособность заемщика, а также скоринговая модель является неотъемлемой частью финансовой системы и имеет большое влияние на кредитный риск.

Несмотря на все преимущества, у модели есть и свои ограничения. В частности, она может быть предвзятой и дискриминировать определенные группы заемщиков. Кроме того, сложность модели может затруднить ее понимание и объяснение.

В будущем планируем улучшить модель, учитывая эти ограничения, также надеемся расширить наше исследование, применяя нашу модель к другим типам кредитных продуктов и рынкам.

СПИСОК ЛИТЕРАТУРЫ

1. Kamimura, E. S., Pinto, A. R. F., Nagano, M. S. (2023). A recent review on optimisation methods applied to Credit Scoring Models. *Journal of Economics, Finance and Administrative Science*1.
2. Bhattacharya, A., Biswas, S. K., Mandal, A. (2023). Credit risk evaluation: a comprehensive study. *Multimedia Tools and Applications*, 82, 18217-182672.
3. World Bank. (n.d.). CREDIT SCORING APPROACHES GUIDELINES3.
4. Hand, D. J. и Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 160(3), 523-541.
5. Lessmann, S., Baesens, B., Seow, H. V. и Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124-136.
6. Yap, B. W., Rani, K. A. и Abdullah, N. N. (2014). An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets. In *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)* (pp. 13-22). Springer, Singapore.

СИНТЕЗ РЕШЕНИЯ НА ОСНОВЕ ЛОГИЧЕСКОГО ВЫВОДА

Герман Ю. О., Герман О. В., Мигалевич С. А.
Факультет компьютерных сетей и систем,
Факультет информационных технологий и управления,
Центр информатизации и инновационных разработок,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: juliagerman@gmail.com

Рассматривается задача синтеза решения проблемы (управления системой) на основе логической модели знаний с временным параметром (тактом). Значение временного шага в общем случае может превосходить 1. Модель не ограничена применением только классической логики. Подход излагается на основе иллюстрации, раскрывающей его суть. Обобщение достигается совмещением логической модели с другими логико-комбинаторными алгебраическими или теоретико-множественными моделями.

ВВЕДЕНИЕ

Задача синтеза управления – это классическая задача теории управления. Алгоритмы ее решения зависят от используемой модели задачи. Использование логических моделей в задачах синтеза мотивируется возросшими возможностями искусственного интеллекта и робототехники [1]. Интерес представляют многообразные варианты задачи логического синтеза, связанные с использованием неклассических логик, комбинированием логических и дискретно-комбинаторных или алгебраических и других моделей [2] и т.п. Ниже рассмотрена задача синтеза управления для логики с временным параметром [3]. Временной параметр в этой модели представляет номер шага (такта). Физическая длительность такта не имеет какого-то специального значения. В [4] представлен вариант синтеза для дизъюнктивных формул с произвольным тактом. В этой работе мы рассмотрим формулы с шагом, равным 1.

I. ПРИМЕР ЗАДАЧИ СИНТЕЗА

Имеется закрытый шлагбаум, автомобиль слева от шлагбаума (позиция A) и площадка справа от шлагбаума (позиция B). Возможны следующие действия:

- **up** (поднять шлагбаум);
- **down** (опустить шлагбаум);
- **move** (переместить автомобиль).

Для описания состояния системы введем следующие предикаты:

- $open(t)$ (шлагбаум открыт в момент (t));
- $closed(t)$ (шлагбаум закрыт в момент (t));
- $A(t)$ (автомобиль на позиции (A) в момент (t));
- $B(t)$ (автомобиль на позиции (B) в момент (t)).

Каждое действие предполагает выполнение предусловий. Само действие k (пред)условиям не относится. Поэтому запишем формулы системы несколько особым образом:

1. **[move]**
 $A(t) \& open(t) \rightarrow B(t+1)$.

Эта формула передает предусловия и постусловие действия переместить автомобиль, записанного в квадратных скобках. Сама логическая формула связывает предусловия и постусловия действия. Далее по аналогии

2. **[up]**
 $closed(t) \rightarrow open(t+1)$.

3. **[down]**
 $open(t) \rightarrow closed(t+1)$.

Кроме этих формул, есть еще «внутренняя» логика системы, не связанная с действиями. Нам нужна такая формула

4. **[]**
 $\neg A(t+1) \rightarrow \neg A(t)$.

Формула 4 означает, что если автомобиля нет на позиции A в момент $t+1$, то его не было там и в момент t . Начальные условия (действий нет)

5. $closed(0)$.
6. $A(0)$.

Сформулируем цель (теорему), которую нужно доказать

7. $\exists t \ closed(t) \& B(t)$.

Выбор обозначений для переменных в целом свободен, если не ведет к нарушению логического смысла). Построение доказательств можно выполнять различными путями. Один из широко используемых методов – метод резолюций Дж. Робинсона. В соответствии с этими правилами логики формулы 1 – 5 примут такой вид:

1. **[move]**
 $\neg A(t) \vee \neg open(t) \vee B(t+1)$.
2. **[up]**
 $open(t+1) \vee \neg closed(t)$.
3. **[down]**
 $closed(t+1) \vee \neg open(t)$.
4. $A(t+1) \vee \neg A(t)$.
5. $closed(0)$.
6. $A(0)$.
7. $\exists w \ closed(w) \& B(w)$.

Согласно технике доказательства от противного, требуется присоединить к системе отрицание доказываемой теоремы и получить вывод пустой формулы (\square) – лжи. Отрицание теоремы имеет такой вид

$$7. \forall w \neg closed(w) \vee \neg B(w)$$

или просто

$$\neg closed(w) \vee \neg B(w).$$

Правило порождения резольвенты поясним на примере дизъюнктов

2. [up]

$$open(t+1) \vee \neg closed(t).$$

3. closed((0)).

Правило получения резольвенты требует сформировать новый дизъюнкт, объединив с помощью операции дизъюнкции унифицированные литералы родительских дизъюнктов и исключив контрарную пару литералов. В итоге получаем такую резольвенту:

$$open(1).$$

В скобках указывается номер такта. Далее, например, из (3) и (7) получаем

8. [down]

$$\neg B(w) \vee open(w-1)$$

(подстановка) $t = w - 1$)

Из (1) и (8) получаем

9. [move, down]

$$\neg A(w-1) \vee \neg open(w-1)$$

(подстановка) $t = w - 1$)

Из (2) и (9) получаем

10. [up, move, down]

$$\neg A(w-1) \vee \neg closed(w-2)$$

(подстановка) $t = w - 2$)

Из (5) и (10) получаем

11. [up, move, down]

$$\neg A(1)$$

(подстановка $w = 2$)

Из (4) и (11) получаем

12. [up, move, down]

$$\neg A(0)$$

(подстановка $t = 0$)

Из (6) и (11) получаем противоречие. Доказательство найдено.

II. ЗАКЛЮЧЕНИЕ

Мы показали, как «привязать» действия к выводу. Каждое действие описывается пред-условиями и постусловиями. При построении резольвент условия обоих участвующих в процессе родительских дизъюнктов соединяются в последовательность. В итоге получена такая финальная последовательность

[up, move, down]

$$\neg A(0)$$

$$A(0)$$

противоречие(\square)

Очевидно, что это общее замечание требует конкретизации в каждом конкретном случае. Сам по себе факт построения плана действий в процессе логического вывода указывает на сильную сторону логического аппарата.

СПИСОК ЛИТЕРАТУРЫ

1. Krishnan, G. A. Artificial intelligence and robotics. / G. Krishnan, M. A. Susmera // SSRN Electronic Journal (www.ijert.org), 6(2). 2018, pp : 1323–1327.
2. Russell, S. Artificial intelligence. A modern approach / S. Russell, P. Norvig // Prentice Hall, 2010, 1192p.
3. Герман, О. В. Метод синтеза поведения интеллектуальной системы / О. В. Герман, О. И. Садовская // Труды БГТУ. Сер. 3, Физ.-мат. науки и информатика. - Минск: БГТУ, 2013. - № 6 (230). - С. 139-141
4. Герман, О. В. Составление плана работ на основе системы дизъюнктов разностного типа / О. В. Герман, Ю. О. Герман, О.И. Садовская // Труды БГТУ. Сер. 3, Физ.-мат. науки и информатика. - Минск: БГТУ, 2014. - № 6 (170). - С. 131-134

МЕТОДИКА РАСЧЕТА СТРУКТУРЫ И ПАРАМЕТРОВ ОБНАРУЖИТЕЛЯ ЛЧМ-СИГНАЛОВ С НЕИЗВЕСТНЫМИ ПАРАМЕТРАМИ

Ковалевич В. В.

Научно-исследовательская лаборатория авиационного факультета Военная академия Республики Беларусь
Минск, Республика Беларусь
E-mail: vlad.kovalevich@gmail.com,

Получено универсальное эмпирическое выражение, описывающее потери выигрыша в отношении сигнал шум (ОСШ) при обработке линейно-частотно модулированного (ЛЧМ) сигнала с неизвестными параметрами на основе использования дискретного преобразования Фурье (ДПФ). Разработана методика обоснования структуры и параметров обнаружителя ЛЧМ-сигналов, обеспечивающего обнаружение сигналов источников радиоизлучения (ИРИ) с потерями выигрыша в ОСШ, не превышающими заданную величину. Произведена оценка эффективности разработанной методики на примере обнаружения ИРИ, используемого ЛЧМ-сигнал. Показано улучшение точности (уменьшение среднеквадратического отклонения (СКО)) пеленгования ИРИ.

ВВЕДЕНИЕ

Современные вооруженные конфликты характеризуются интенсивным применением противоборствующими сторонами радиоэлектронных средств (РЭС) различного назначения. Обнаружение излучений РЭС на большем удалении, позволяет принять своевременные меры по их огневому поражению и радиоэлектронному подавлению. При этом особую актуальность имеет задача пеленгования и определения местоположения РЭС применяющие ЛЧМ-сигналы.

В настоящее время разработано достаточно большое количество методов обнаружения и оценки параметров ЛЧМ – сигналов с неизвестными параметрами [1, 3]. Всем разработанным методам в большей или меньшей степени свойственно, для обнаружения и оценки параметров ЛЧМ-сигнала построение трехмерной зависимости и поиск ее максимума [1]. Применение ДПФ, как одного из возможных способов обработки ЛЧМ-сигнала, позволяет это выполнять [2].

Способ обработки ЛЧМ-сигнала основан на его особенностях и обеспечении взаимосвязи параметров устройства обработки со значением крутизны закона ЛЧМ, что обуславливает его универсальность для различных сочетаний частотно-временных параметров, характеризующихся одним и тем же значением скорости изменения частоты сигнала [2]. Согласование параметров обнаружителя со скоростью изменения частоты ЛЧМ-сигнала обеспечивает «квазиоптимальное» обнаружение сигналов ИРИ. Под термином «квазиоптимальное» обнаружение ИРИ будем понимать обнаружение сигналов с наилучшим выигрышем (W) в ОСШ на выходе устройства обработки (q_{out}^2) в условиях априорной неопределенности относительно вида и параметров закона модуляции сигналов ИРИ. «Квазиоптимальное» обнаружение ИРИ, обеспечивается за счет сужения эквивалентной шумовой полосы пропускания фильтров ДПФ до значения, определяе-

мого скоростью изменения частоты (крутизной) ЛЧМ-сигнала. Использование «квазиоптимального» устройства обнаружения позволяет получить выигрыш в ОСШ, равный корню квадратному из базы ЛЧМ-сигнала [2]

$$q_{out}^2 = \frac{E_c}{N_0} = \frac{P_c \tau}{P_n / \Delta F_M} = q_{in}^2 \sqrt{B}. \quad (1)$$

где q_{in}^2, q_{out}^2 – отношение сигнал-шум по мощности на входе и выходе «квазиоптимального» устройства обнаружения; E_c – энергия принятого сигнала; N_0 – спектральная плотность мощности входного шума; P_c, P_n – импульсная мощность полезного сигнала и средняя мощность шума на входе «квазиоптимального» устройства обнаружения; $\tau, \Delta F_M, B$ – длительность импульса, девиация и база ЛЧМ-сигнала.

I. ПОСТАНОВКА ЗАДАЧИ

Пусть известны диапазоны возможных значений девиации, длительности импульса и несущей частоты ЛЧМ – сигнала ИРИ.

Требуется обосновать методику расчета структуры и параметров обнаружителя, обеспечивающего «квазиоптимальное» обнаружение ЛЧМ – сигнала ИРИ. Расчет структуры и параметров обнаружителя является двухкритериальной оптимизационной задачей определения минимально необходимого количества каналов обнаружения при максимальном значении выигрыша в отношении сигнал/шум (минимальных потерях выигрыша) в заданном диапазоне изменения значений девиации, длительности импульса и несущей частоты ЛЧМ – сигнала.

II. РЕШЕНИЕ

Для разработки методики расчета параметров «квазиоптимального» обнаружителя ЛЧМ-сигналов с применением ДПФ требуется аналитическое выражение, описывающее снижение выигрыша в ОСШ. На основе результатов ИММ [4] были получено эмпирическое выражение, описы-

вающие потери выигрыша в ОСШ.

$$L = \frac{2\gamma}{\gamma_c} \left(0.5 + \frac{\gamma}{\gamma_c} \right)^{-2}. \quad (2)$$

где $\gamma_c \in (\gamma_{cmin} \dots \gamma_{cmax})$ – диапазон изменения значений крутизны ЛЧМ-сигнала; $\gamma = \Delta f_{DFT}^2$.

Методика расчета структуры и параметров «квазиоптимального» обнаружителя ЛЧМ-сигналов должна обеспечивать расчет количества каналов ДПФ, для перекрытия возможного диапазона крутизны ЛЧМ-сигнала с потерями выигрыша в ОСШ не более заданных, а также получение количества отсчетов для выполнения ДПФ.

Методика расчета структуры и параметров «квазиоптимального» обнаружителя ЛЧМ-сигналов, заключается в определении минимально достаточного количества каналов ДПФ и номинальных значений их полос пропускания (длительности выборки или количества отсчетов ДПФ) для обеспечения обработки ЛЧМ-сигнала с потерями, не более заданных (L_{max}), при произвольном значении крутизны ЛЧМ-сигнала.

Методика обоснования структуры и параметров «квазиоптимального» обнаружителя ЛЧМ-сигналов описывается следующей последовательностью операций:

1. Ввод возможных параметров ЛЧМ-сигнала и уровень допустимых потерь (L_{max}).
2. Оценка требуемого количество каналов ДПФ для перекрытия диапазона возможной крутизны ЛЧМ-сигнала

$$N_{DPF} = \frac{\lg(\Delta F_{Mmax}\tau_{max}/\Delta F_{Mmin}\tau_{min})}{\lg\left(4\left(\frac{1}{L_{max}} + \sqrt{\frac{1}{L_{max}^2} + \frac{1}{L_{max}}} - \frac{1}{2}\right)^2\right)}. \quad (3)$$

3. Определение значений «квазиоптимальной» крутизны, по формуле

$$\gamma_k = \frac{1}{2}\gamma_{min} 10^{\frac{2k-1}{2N_{DPF}} \lg \frac{\Delta F_{Mmax}\tau_{max}}{\Delta F_{Mmin}\tau_{min}}}, \quad (4)$$

где $k = 1, 2, \dots, N_{DPF}$.

4. Расчет требуемого количества отсчетов при выполнении ДПФ с учетом частоты поступления отсчетов ЛЧМ-сигнала на вход «квазиоптимального» обнаружителя

$$N_{points} = \frac{F_s}{\sqrt{\gamma_k}}. \quad (5)$$

Рассмотрим результаты применения предложенной методики для обоснования структуры и параметров «квазиоптимального» обнаружителя ЛЧМ-сигналов зенитного ракетного комплекса (ЗРК) типа «PATRIOT», с девиацией и длительностью импульсов примерно равными $\Delta F_M = 1 \dots 10$ МГц и $\tau = 10 \dots 100$ мкс. При этом максимально допустимые потери выигрыша в ОСШ задано 1 дБ ($L_{max} = 1$), а частота дискретизации ($F_s = 100$ МГц). Зависимости потерь выигрыша в ОСШ на выходе канала ДПФ с полученными значениями

пороговой и оптимальной крутизны представлены на рисунке 1.

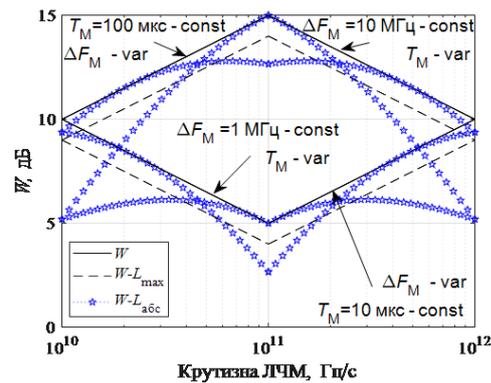


Рис. 1 – Графики зависимости выигрыша в отношении сигнал/шум от крутизны ЛЧМ сигнала при заданных максимально допустимых потерях

Как видно из рисунка 1 в результате обоснования структуры обнаружителя получено симметричное расположение каналов ДПФ, а их необходимое количество ровно укладывается в заданный диапазон возможной крутизны ЛЧМ – сигнала. Зависимость потерь в каналах не превышает заданных и даже не доходят до них.

Выигрыш в ОСШ полученный за счет применения «квазиоптимального» обнаружения позволяет увеличить дальность обнаружения и точность пеленгования (уменьшить СКО) в корень квадратный из выигрыша в раз, для рассмотренного примера составит 1,78 – 5,62 раза.

Таким образом, выполнено аналитическое обоснование структуры и параметров «квазиоптимального» обнаружителя зондирующих ЛЧМ-сигналов, обеспечивающей обнаружение сигналов с потерями выигрыша в отношении сигнал-шум, не превышающих заданные.

Полученные количественные характеристики выигрыша в энергетике ЛЧМ-сигнала позволяют говорить о возможности существенного расширения зон действия аппаратуры РТР и повышения точности пеленгования ИРИ при сохранении и уменьшении массогабаритных характеристик аппаратуры.

СПИСОК ЛИТЕРАТУРЫ

1. Канатчиков А.А. Сравнительный анализ методов обнаружения и определения параметров зондирующих сигналов РСА космического базирования. Журнал радиоэлектроники [электронный журнал]. 2020. №7.
2. Гатальский А. М. / Обоснование способа цифровой обработки ЛЧМ-сигнала на основе применения дискретного преобразования Фурье / А. М. Гатальский, А. Г. Боровой вестник ВА №2 2021
3. Методы и средства радиоэлектронной защиты : учеб. пособие / С. В. Козлов. – Минск : БГУИР, 2019. – 188 с. : ил. ISBN 978-985-543-501-4
4. Моделирование системы обнаружения широкополосных сигналов радиовысотометров и радиолокационных взрывателей (шифр «Ртуть»): отчет о НИР (заключ.) / Воен. Акад. Респ. Беларусь; рук. А. М. Гатальский ; исполн.: В. В. Ковалевич [и др.]. – Минск, 2021. – 92 с.

МОДЕЛИРОВАНИЕ ДВИЖЕНИЯ ЛАЗЕРА ПО ЗАДАННОЙ ТРАЕКТОРИИ ДЛЯ ПРОЦЕССОВ ФОРМИРОВАНИЯ ПЕРЕХОДНЫХ ОТВЕРСТИЙ

Лаппо А. И., Ярмолик В. И., Муравьев А. С.

Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: lappo@gmail.com

В данной статье рассматривается проблема формирования переходных отверстий в кремниевой пластине с использованием лазерного луча движущегося по заданной траектории

ВВЕДЕНИЕ

Использование лазерного излучения для технологических операций без контактного воздействия на материал обладает рядом преимуществ, таких как точность следования технологическим режимам и контроль объема подводимой энергии.

Для создания переходных отверстий между слоями многокристального модуля можно применять энергию сфокусированного лазерного луча. Так как диаметр отверстий может быть больше диаметра лазерного луча, приходится использовать способы перемещения луча по заданной траектории. Перемещение луча по траектории может осуществляться способом перемещения предметно-стола относительно неподвижного лазерного луча, так и способом отклонения луча с помощью системы зеркал.

В данном случае, в технологической операции создания переходных отверстий в кремниевой подложке можно выделить две проблемы: первая – выбор технологических режимов работы лазера (мощность, частота импульсов, скорость перемещения лазерного пучка), вторая – выбор оптимальной траектории перемещения луча. Технология формирования точечного отверстия несовершенна, поэтому получаемые точечные отверстия имеют конусообразность, которая определяется как отношение разности диаметров отверстий к толщине подложки. Для качественного отверстия конусообразность не должна превышать 0,1. Учитывая данный факт, формирование отверстий с применением траектории движения лазерного луча по контуру окружности является нецелесообразным, и встает задача проектирования траектории движения с заполнением площади ограниченной окружностью.

I. ВЫБОР ТЕХНОЛОГИЧЕСКИХ РЕЖИМОВ

Для операции создания переходных отверстий в кремниевой подложке планируется использовать лазерное оборудование со следующими характеристиками:

- лазер мощностью 10 Вт,
- длительность импульса – 1,2 нс,
- частота – 150 Гц,
- скважность импульса – 0,67,

– скорость луча – 0,5 или 5 мм/с.

Для действующей длины волны лазера предполагается, что поглощающая способность равна излучающей способности. Следовательно, все лазерное излучение будет уходить на нагрев пластины. Диаметр формируемого отверстия 0,1 мм. Расчет лазерного излучения при формировании переходных отверстий приведен в источнике [1,2].

II. ВЫБОР ВАРИАНТОВ ДВИЖЕНИЯ ПО ЗАДАННОЙ ТРАЕКТОРИИ

Так как лазерное излучение является импульсным, следовательно, по ходу движения луча будут образовываться области, на которых не будет происходить воздействие на материал. Для моделирования траектории движения лазера предложено несколько вариантов, один из них – это движение луча по гипотрохоиде (рис. 1), второй – движение по спирали (рис. 2).

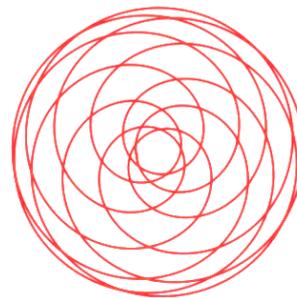


Рис. 1 – Визуализация движения лазера при формировании отверстия по гипотрохоидной траектории

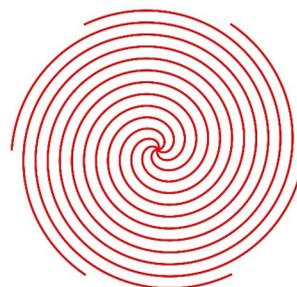


Рис. 2 – Визуализация движения лазера при формировании отверстия по спиральной траектории

Формула движение луча по гипотрохоиде выглядит следующим образом:

$$\begin{cases} x = (r - mr) \cos(mt) + \cos(t - mt) \\ y = (r - mr) \sin(mt) + \sin(t - mt) \end{cases}$$

где: t – время; x, y – координаты; r, m, h – коэффициенты.

Формула движение луча по спирали приведена ниже:

$$\begin{cases} x = \pi \cos(\theta) + at \cos(t + q) \\ y = \pi \sin(\theta) + at \sin(t + q) \end{cases}$$

Для каждого контура получено два варианта траектории движения, в зависимости от скорости движения луча 0,5 или 5 мм/с. (рис. 3 – 6).

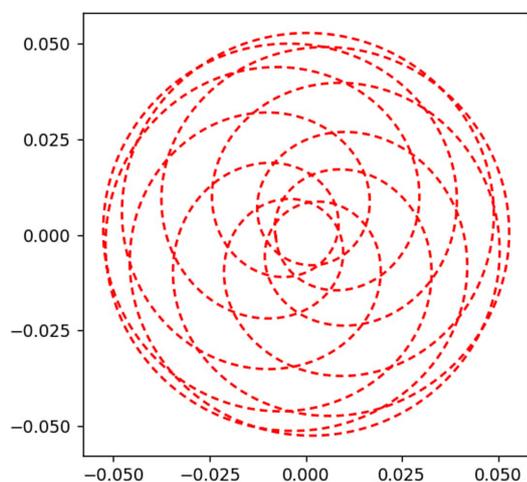


Рис. 3 – Визуализация движения лазера по гипотрохоидной траектории скорости движения луча 0,5 мм/с.

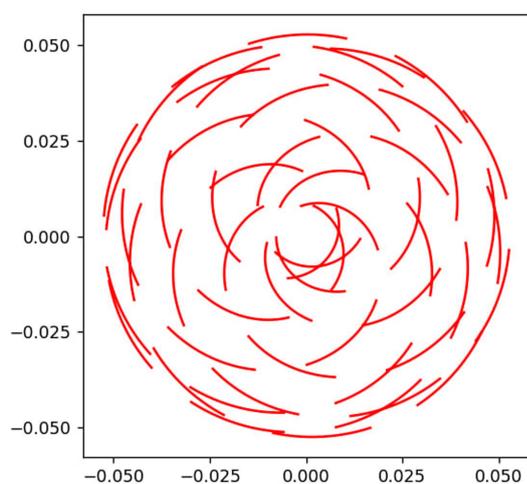


Рис. 4 – Визуализация движения лазера по гипотрохоидной траектории скорости движения луча 5 мм/с.

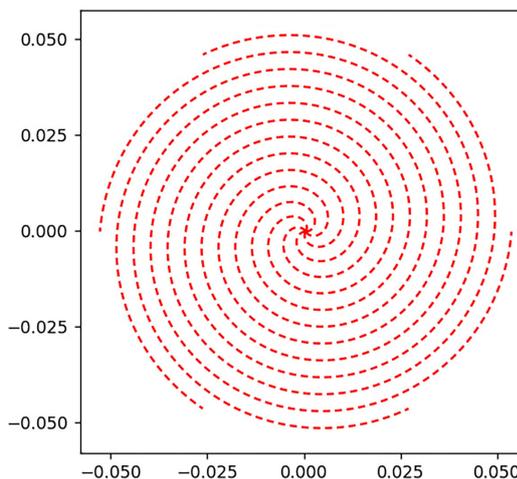


Рис. 5 – Визуализация движения лазера по спиральной траектории скорости движения луча 0,5 мм/с.

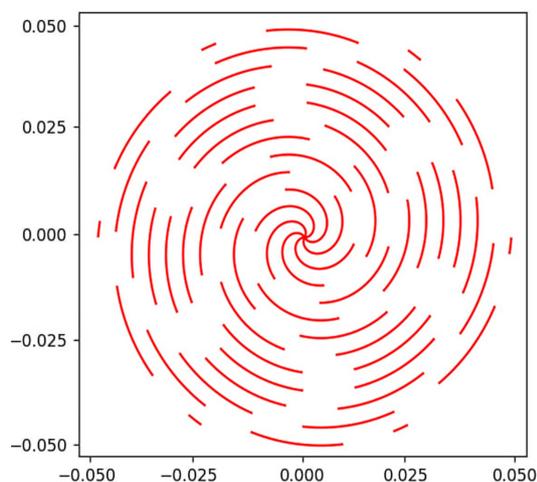


Рис. 6 – Визуализация движения лазера по спиральной траектории скорости движения луча 5 мм/с.

III. Выводы

Таким образом, при формировании отверстий с использованием траектории движения луча по гипотрохоиде со скоростью 0,5 мм/с потребовалось совершить минимум 4 прохода. При использовании движения луча со скоростью 5 мм/с оказалось достаточно 2 проходов по траектории. При спиральном движении луча не получилось создать замкнутый контур.

Стоит отметить равномерное распределение лучей по всему участку обработки материала при использовании движения со скоростью 0,5 мм/с.

1. Лаппо, А.И. Моделирование процесса лазерной прошивки отверстий в кремнии при формировании 3D структур / А.И. Лаппо, Т.С. Боброва, О.В. Кузнецова / Материалы международной научной конференции «ITS-2019» Минск: БГУИР, 2019. – С. 232-233
2. Ланин, В.Л. Формирование отверстий в кремниевой подложке 3D электронного модуля лазерным излучением / В.Л. Ланин, В.Т. Фам, А.И. Лаппо // Доклады БГУИР, 2021, №3. – С. 58-65.

ОЦЕНКА ЭФФЕКТИВНОСТИ ВНЕДРЕНИЯ НОВОГО ПОКОЛЕНИЯ ОБРАЗОВАТЕЛЬНЫХ СТАНДАРТОВ

Литвинович А. И., Горячко А. Д.

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

E-mail: sashka.litvinovich@gmail.com, harachka.aliaksei@gmail.com

Современное высшее образование ставит перед собой задачу обеспечения качественного обучения и формирования конкурентоспособных специалистов. Целенаправленное улучшение образовательного процесса и обеспечение высокого уровня академических знаний студентов являются ключевыми вопросами, которые затрагивают многие аспекты образовательной практики.

ВВЕДЕНИЕ

В работе проводится анализ результатов внедрения нового поколения образовательных стандартов на основе результатов сдачи семестровых экзаменов. Для расчетов были использованы экзаменационные отметки студентов специальности «Автоматизированные системы обработки информации».

I. АНАЛИЗ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО СТАРЫМ ОБРАЗОВАТЕЛЬНЫМ СТАНДАРТАМ

В исследовании использовались отметки студентов специальности АСОИ, обучавшихся по программам разных поколений стандартов (в дальнейшем новое поколение и старое поколение). Для оценки старого поколения стандартов использовались экзаменационные отметки студентов групп 020601-605, а для оценки нового поколения стандартов – экзаменационные отметки студентов групп 120601-605. Целью исследования является оценка эффективности внедрения нового поколения стандартов специальности, выявление проблем и путей их решения. В работе рассчитывались такие показатели, как средний балл студентов по предметам, медиана, мода и дисперсия отметок. Информация о среднем балле по предметам старого поколения стандартов отображена в таблице 1.

Таблица 1 – Средние баллы обучающихся по дисциплинам старого поколения стандартов

Статистические показатели	№	№	№	№	№
	се-мест-ра	се-мест-ра	се-мест-ра	се-мест-ра	се-мест-ра
	1	2	3	4	5
Средний балл	6.21	6.72	6.70	7.08	7.74
Медиана	6.8	6.14	6.36	7.0	7.70
Дисперсия	1.35	2.24	2.39	2.61	3.17

Высокая дисперсия отметок говорит о неоднородном процессе обучения и необходимости улучшения качества образования. Для оценки академического прогресса студентов, рассчитана частота вхождения отметок в заданный диапазон (рисунок 1).



Рис. 1 – Частота вхождения студентов в определенный диапазон

Высокая дисперсия отметок говорит о неоднородном процессе обучения и необходимости улучшения качества образования.

II. АНАЛИЗ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО НОВЫМ ОБРАЗОВАТЕЛЬНЫМ СТАНДАРТАМ

Внедрение новых образовательных стандартов позволило расширить использование современных информационных технологий и новых программных продуктов в процессе обучения что должно было улучшить качество обучения (таблица 2).

Таблица 2 – Средние баллы обучающихся по дисциплинам старого поколения стандартов

Статистические показатели	№ се-местра	№ се-местра	№ се-местра
	1	2	3
Средний балл	7.27	7.28	7.48
Медиана	7.2	7.375	7.44
Дисперсия	1.37	3.48	1.37

Частота вхождения отметок в заданный диапазон представлена на рисунке 2.



Рис. 2 – Частота вхождения студентов потока 12060х в определенный диапазон

III. ЗАКЛЮЧЕНИЕ

В результате проведенных исследований установлено, что средний балл потока студентов повысится при переходе на обучение по новому образовательному стандарту. Величина рассчитанной медианы отметок говорит о наличии студентов, у которых имеются существенные пробелы в знаниях по некоторым дисциплинам. На основе проведенного анализа отметок студентов можно сделать вывод что переход на новый образовательный стандарт привел к улучшению качества обучения студентов.

СТАТИСТИЧЕСКАЯ КЛАССИФИКАЦИЯ СТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ РАЗЛОЖЕНИЯ ВОЛЬДА

Микулич Г. В., Жук Е. Е.

Кафедра математического моделирования и анализа данных,

Белорусский государственный университет

Минск, Республика Беларусь

E-mail: aragornuga@gmail.com, zhukkee@mail.ru

Рассматривается проблема статистической классификации реализаций стационарных в широком смысле временных рядов, различающихся по классам значениями своих параметров – коэффициентов авторегрессии из разложения Вольда.

ВВЕДЕНИЕ

Задача статистической классификации является одной из основных прикладных задач математической статистики. При этом, данные, подлежащие классификации, во многих случаях находятся в формате временных рядов, например: биомедицинские измерения (электрокардиограмма, кровяное давление), данные о погоде, цены акций на бирже и др. В данной работе изучается случай стационарных временных рядов, а в качестве процедуры классификации выбран кластерный анализ.

I. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Пусть наблюдается случайная выборка $X^n = (X_1, \dots, X_n)$ объёма n из независимых в совокупности случайных векторов наблюдений, принадлежащих к $L \geq 2$ классам $\Omega_1, \dots, \Omega_L$. Наблюдение X_t принадлежит к классу со случайным ненаблюдаемым номером $d_t^0 \in S$, $S = \{1, \dots, L\}$, $t = \overline{1, n}$ и при фиксированном номере класса $d_t^0 = i$, $i \in S$ является реализацией длительности T_t ($X_t = (x_{t1} \dots, x_{tT_t})' \in R^{T_t}$, ' – символ транспонирования) временного ряда авторегрессии $x^i = \{x_{tj}^i\}_{j=-\infty}^{+\infty}$ порядка $p \geq 1$ (модель AP(p))

$$x_l^i + \theta_{i1}^0 x_{l-1}^i + \dots + \theta_{ip}^0 x_{l-p}^i = u_l^i, \quad l \in Z, \quad (1)$$

где $Z = \{0, \pm 1, \pm 2, \dots\}$, $\theta_{ij}^0 \in R^p$ – вектор авторегрессии для i -го класса, а $\{u_l^i\}_{l=-\infty}^{+\infty}$ – независимые в совокупности нормальные случайные величины с нулевым математическим ожиданием и одинаковой дисперсией σ^2 для всех классов Ω_i :

$$E \{u_l^i\} = 0, \quad D \{u_l^i\} = \sigma^2, \quad l \in Z, \quad i \in S. \quad (2)$$

Наряду с вектором коэффициентов авторегрессии θ_i^0 класс Ω_i также характеризуется своей априорной вероятностью:

$$P \{d_t^0 = i\} = \pi_i^0 > 0, \quad i \in S, \quad \sum_{i=1}^L \pi_i^0 = 1. \quad (3)$$

II. ПРЕДВАРИТЕЛЬНЫЕ ПРЕОБРАЗОВАНИЯ

Пусть имеется стационарный в широком смысле временной ряд. В модели (1) – (3) порядок авторегрессии полагается известным, но в нашем случае приходится дополнительно найти для него оценку \hat{p} . Эта задача сводится с помощью разложения Вольда

$$x_l + \sum_{j=1}^{\infty} \theta_j x_{l-j} = u_l, \quad l \in Z$$

к оцениванию неизвестных коэффициентов авторегрессии θ_j (u_l удовлетворяет условиям (2)). Пошаговая процедура оценивания описана в [1].

Вернёмся к модели (1) – (3). Теперь будем полагать порядок авторегрессии p известным и равным полученной оценке \hat{p} . Пусть для коэффициентов авторегрессии выполняется следующее соотношение:

$$z^p + \sum_{j=1}^p \theta_{ij} z^{p-j} = 0 \quad (4)$$

Преобразуем исходную выборку X^n в выборку $Y^n = (Y_1, \dots, Y_n)$, где $Y_t \in R^p$, $t = \overline{1, n}$ – МП-оценка для p -вектора коэффициентов авторегрессии $\theta_{d_t^0}^0 \in R^p$, построенная по наблюдению $X_t \in R^{T_t}$, являющемуся реализацией длительности T_t одного из временных рядов AP(p) из (1).

Для построения МП-оценки Y_t воспользуемся тем фактом, что наблюдение $X_t \in R^{T_t}$ при фиксированном $d_t^0 = i$, $i \in S$ является нормальным T_t -вектором с нулевым математическим ожиданием $E \{X_t | d_t^0 = i\} = 0_{T_t}$, и плотностью распределения вероятностей (5), указанной в конце документа.

В формуле (5) $n_p(y|\mu, \Sigma)$ – плотность p -мерного нормального распределения, $X_t^p = (x_{t1}, \dots, x_{tp})' \in R^p$, $R_p(\theta_i^0, \sigma) = E \{X_t^p (X_t^p)' | d_t^0 = i\}$ – невырожденная [2] ковариационная матрица, элементами которой являются автоковариации $(\rho_{|k-l|}(\theta_i^0, \sigma))_{k,l=1}^p$, определяемые системой уравне-

ний Юла – Уокера [3]:

$$\sum_{j=1}^p \theta_{ij}^0 \rho_j(\theta_i^0, \sigma) = \sigma^2;$$

$$\sum_{j=1}^p \theta_{ij}^0 \rho_{|k-j|}(\theta_i^0, \sigma) + \rho_k(\theta_i^0, \sigma) = 0, \quad k = 1, 2, \dots$$

Согласно методу максимального правдоподобия,

$$\{Y_t, \hat{\sigma}_t\} = \arg \max_{\{\bar{\theta}, \sigma\}} \ln p(X_t; \bar{\theta}, \sigma), \quad (6)$$

где $p(X_t; \bar{\theta}, \sigma)$ – плотность из (5), записанная для $\bar{\theta}$, $\theta_i^0 := \bar{\theta}$. Отметим, что при решении задачи (6) наряду с Y_t строится также оценка $\hat{\sigma}_t^2$ для дисперсии.

III. ИССЛЕДОВАНИЕ ПОЛУЧЕННЫХ ОЦЕНОК

Теорема. Пусть в условиях модели (1) – (3) корни характеристических уравнений (4) лежат внутри единичного круга. Тогда при фиксированном векторе истинной классификации $D^0 = (d_1^0, \dots, d_n^0)$ МП-оценки Y_1, \dots, Y_n , полученные из (6) являются состоятельными оценками для соответствующих параметров авторегрессии:

$$Y_t \xrightarrow{P} \theta_{d_t^0}^0, \quad T_t \rightarrow +\infty,$$

и имеют асимптотически нормальное распределение

$$\sqrt{T_t}(Y_t - \theta_{d_t^0}^0) \rightsquigarrow N_p(0_p, \sigma^2 R_p^{-1}(\theta_{d_t^0}^0, \sigma)), \quad T_t \rightarrow +\infty.$$

Доказательство. Справедливость полученных соотношений следует из статистических свойств МП-оценок коэффициентов авторегрессии [1] и того факта, что при фиксированном $d_t^0 = i$, $i \in S$ исходное наблюдение X_t является реализацией длительности T_t временного ряда $AR(p)$ с вектором коэффициентов авторегрессии $\theta_{d_t^0}^0$, а Y_t – МП-оценкой для $\theta_{d_t^0}^0$ из (6), полученной по X_t , $t = \overline{1, n}$.

IV. КЛАСТЕР-ПРОЦЕДУРА

Построим процедуру кластер-анализа в пространстве МП-оценок параметров авторегрессии, основанную на алгоритме L -средних.

1. По исходной выборке X^n из (6) находится выборка МП-оценок Y^n , из которой в качестве начальных приближений $\{\hat{\theta}_i^{(0)}\}$, $i \in S$ для «центров» $\{\theta_i^0\}$, $i \in S$ классов $\{\Omega_i\}$ выбираются какие-либо L наблюдений Y_{j_1}, \dots, Y_{j_L} , $j_i \in \{1, \dots, n\}$; $j_i \neq j_k$, $i \neq k \in S$;

2. На l -м шаге ($l = 0, 1, 2, \dots$) производится классификация выборки Y^n :

$$(\hat{d}_t^{(l)}) = \arg \min_{i \in S} |Y_t - \hat{\theta}_i^{(l-1)}|, \quad t = \overline{1, n},$$

т.е. строится оценка $\hat{D}^{(l)} = (\hat{d}_1^{(l)}, \dots, \hat{d}_n^{(l)})'$ для D^0 , и уточняются оценки для «центров» классов:

$$\hat{\theta}_i^{(l)} = \left(\sum_{t=1}^n \delta_{\hat{d}_t^{(l)}, i} \right)^{-1} \sum_{t=1}^n \delta_{\hat{d}_t^{(l)}, i} Y_t, \quad i \in S$$

где $\delta_{i,j}$ – символ Кронекера;

3. Итерационный процесс останавливается при достижении на l -м шаге ($2 \leq l < \infty$) равенства $\hat{D}^{(l)} = \hat{D}^{(l-1)}$, и его результатом являются оценки $\hat{D} := \hat{D}^{(l)}$ для вектора истинной классификации D^0 и $\hat{\theta} := \hat{\theta}^{(l)} \in R^{Lp}$ для составного вектора θ^0 параметров авторегрессии θ_i^0 , $i \in S$.

Замечание. Как показано в [2], при больших длительностях наблюдений ($T_t \rightarrow +\infty$) из исходной выборки в приведённом выше алгоритме вместо МП-оценок можно использовать МНК-оценки:

$$Y_t = - \left(\sum_{l=p+1}^{T_t} X_{tl}^p (X_{tl}^p)' \right)^{-1} \sum_{l=p+1}^{T_t} x_{tl} (X_{tl}^p)',$$

$$X_{tl}^p = (x_{t,l-1}, \dots, x_{t,l-p})' \in R^p, \quad t = \overline{1, n},$$

или оценки Юла – Уокера:

$$Y_t = -(\hat{R}_t^p)^{-1} \hat{r}_t^p, \quad \hat{R}_t^p = (\hat{\rho}_{|k-l|}^t)_{k,l=1}^p,$$

$$\hat{r}_t^p = (\hat{\rho}_1^t, \dots, \hat{\rho}_p^t)', \quad t = \overline{1, n}.$$

V. ЗАКЛЮЧЕНИЕ

Рассмотрена проблема классификации стационарных временных рядов. Предложен алгоритм кластерного анализа в пространстве оценок максимального правдоподобия параметров авторегрессии, основанный на алгоритме L -средних.

1. Харин Ю. .С. Теория вероятностей, математическая и прикладная статистика : учеб. пособие / Ю. С. Харин, Н. М. Зуев, Е. Е. Жук. – Минск: БГУ, 2011. – 464 с.
2. Бокс Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс. – М. : Мир, 1974. – 406 с.
3. Андерсон Т. Статистический анализ временных рядов / Т. Андерсон. – М. : Мир, 1976. – 760 с.

$$p(X_t; \theta_i^0, \sigma) = n_p(X_t^p | 0_p, R_p(\theta_i^0, \sigma)) \times$$

$$\times (2\pi)^{-\frac{T_t-p}{2}} \sigma^{-(T_t-p)} \exp \left(-\frac{1}{2\sigma^2} \sum_{l=p+1}^{T_t} (x_{tl} + \theta_{i_1}^0 x_{t,l-1} + \dots + \theta_{i_p}^0 x_{t,l-p})^2 \right) \quad (5)$$

ЭНЕРГОЭФФЕКТИВНЫЙ АЛГОРИТМ АВТОМАТИЧЕСКОЙ ЗАМЕНЫ ФОНА ДЛЯ ВИДЕОКОНФЕРЕНЦИЙ

Пирштук Д. И.

Кафедра дискретной математики и алгоритмики, Факультет прикладной математики и информатики,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: PirshukDI@bsu.by

В работе предложен алгоритм отделения человека от фона на видео на основе быстрой полносверточной сегментационной нейронной сети и специальной процедуры обнаружения движения сегментируемого объекта в кадре. Предлагаемый алгоритм потребляет около 3-6% CPU при обработке HD-видеопотока скоростью 30 кадров/сек и может использоваться для замены виртуального фона в видеоконференциях и других случаях дополненной реальности без существенного увеличения энергопотребления, разрядки аккумулятора ноутбука или мобильного телефона.

ВВЕДЕНИЕ

Для эффективной удаленной коммуникации между людьми важны не только передача аудио, но и видео, так как большинство людей являются визуалами, которые в процессе диалога любят использовать жесты и хотят видеть эмоции и жесты собеседников. Поэтому организаторы видеоконференций настоятельно просят участников включать камеру во время созвонов. Но с другой стороны включенная камера при созвоне из дома создает конфликт между рабочим и личным пространством.

Замена фона с использованием высококонтрастного однотонного экрана (хромакея) хоть и является хорошим аппаратным решением, но очевидно неудобно для домашнего использования. Поэтому является актуальной задача программно отделить человека от фона в обычном видеопотоке с веб-камеры и заменить фон на виртуальный. Обработка видеопотока в реальном времени в облаке затруднительна из-за затрат, высоких задержек при передаче данных через Интернет и проблем конфиденциальности данных. Решение должно работать именно на клиентском устройстве, причем функционал виртуальной замены фона во время видеозвонка должен потреблять минимальное количество дополнительных вычислительных ресурсов, так как прием-передача и отображение нескольких видеопотоков уже и так является достаточно ресурсоемкой задачей. Более того, в отличие от настольных компьютеров, многие ноутбуки и смартфоны не могут долго работать с высокой производительностью. Либо процессор перегревается и снижает тактовую частоту, либо быстро разряжается аккумулятор.

Предлагаемое решение состоит из двух компонент – модуля обнаружения движений в кадре и нейронной сети для очень быстрой сегментации человека в кадре, и потребляет около 3-6% CPU.

I. БЫСТРАЯ СЕГМЕНТАЦИЯ ЧЕЛОВЕКА

Для отделения человека от фона предлагается использовать асимметричную полносверточную U-Net-подобную [1] архитектуру типа Encoder-Decoder с использованием MobileNetV3Small-0.75x [2] в качестве кодировщика и легковесным декодером. Для объединения признаков использовалась модификация модуля Squeeze-and-excitation [3] в качестве модуля внимания. В качестве функции потерь использовалась функция потерь, описанная в [4].

Первоначальное обучение нейросети выполнялось в разрешении 256×256 с последующим дообучением под рабочее разрешение 256×144 . Заметим, что это ровно $1/5$ от HD-разрешения 1280×720 , наиболее часто используемого в видеозвонках, что позволяет эффективно уменьшать размер кадра без дополнительной интерполяции.

Предварительное обучение выполнялось на смеси разных изображений: фотографии людей в различных условиях в полный и частичный рост со всевозможными перекрытиями и другими объектами в кадре, профессиональные портреты, селфи, снятые на камеру мобильного телефона, все возможные интерьеры без человека и др. Но заметим, что сверточные сети не инвариантны к изменению масштаба объектов и перекрытиям. Частично эта проблема может быть компенсирована увеличением количества фильтров на слоях и модулем контекстной агрегации. Для достижения максимального качества сегментации при минимальном количестве вычислений вместо этого мы выполнили дообучение на размеченных кадрах видеозвонков для адаптации сети под сегментацию человека, сидящего чаще всего за столом при различных освещениях и жестикулирующего.

Для оценки качества использовалась метрика Intersection-over-Union (IoU) между размеченной и предсказанной масками $\frac{TP}{TP+FP+FN}$, где TP , FP и FN – количество истинно-положительных, ложно-положительных и ложно-отрицательных пикселей соответственно.

Предложенная архитектура требует около 59 миллионов операций с плавающей точкой для обработки одного кадра, что занимает около 4 мс на обработку одного кадра в один поток на ноутбуке с процессором Intel Core i7-7500U на фреймворке Tensorflow Lite с использованием XNNPACK делегата [5]. Обученная нейросеть достигла точности сегментации по метрике IoU в 96,04% на нашем тестовом датасете. Пример сегментации на кадре из тестового видео приведен на рисунке 1.

II. ОПРЕДЕЛЕНИЕ ДВИЖЕНИЯ В КАДРЕ

Заметим, что движения рта, изменения выражения лица, движения глаз не влияют на границу между человеком и фоном, а изменятся граница в следующих случаях:

1. Выход части тела за пределы кадра.
2. Появление части тела в кадре.
3. Движение частей тела человека, например, жесты или кивки головой.

Предлагается следующая процедура:

1. Выделим все граничные пиксели между человеком и фоном путем усредняющего размытия предсказанной сегментационной маски с ядром 11 и включим в границу пиксели со значениями маски 0,2–0,8. Количество пикселей на границе обозначим через N .
2. Возьмем в качестве области интереса все пиксели границы, а также полоски по 5 пикселя периметра кадра.
3. Будем считать, что пиксель нового кадра подозрительный, если он принадлежит области интереса и изменился в оттенках серого на 10% диапазона 0–255. Пусть M – количество подозрительных пикселей.
4. Не пересчитываем сегментационную маску, если $M/(N + 1) < 2,5\%$ и количество непрерывно пропущенных кадров меньше 30.

Процедура применяется к уменьшенному до разрешения 256×144 изображению. Константы подобраны по датасету видеозвонков и обусловлены в том числе тем, что веб-камеры большинства потребительских устройств достаточно сильно шумят.

Эксперименты показывают, что люди удовлетворены качеством замены фона, если IoU больше 95%. Алгоритм допускает пропуск пересчета сегментационной маски, если метрика IoU

для нового кадра со старой маской не может быть более, чем на 2,5% хуже IoU для изначальной пары. Так как среднее IoU нашей сегментационной нейронной сети больше 96%, то если сегментация на исходной паре имела высокое значение метрики IoU, то и для нового кадра со старой маской при выполнении критерия пропуска значение IoU будет удовлетворительным. При этом на случай, когда изначальная пара была случайно отсеgmentирована сильно хуже среднего ожидаемого IoU и отслеживаемая граница неверная, предусмотрен принудительный пересчет хотя бы раз в 30 кадров.

Данная процедура позволяет пропускать на видеозвонках в среднем около 60% кадров, снижая среднее потребление CPU до 3–6%.

ЗАКЛЮЧЕНИЕ

Предложен алгоритм отделения человека от фона на видео на основе нейронной сети и критерия обновления сегментационной маски со средним потреблением на обычных CPU потребительских устройств в 3–6% при обработке видеопотока 30 кадров/сек. Предложение может быть также использовано для других задач улучшения видеозвонков, таких как сглаживание кожи, перекраска губ, перекраска волос, и в других сегментационных задачах обработки видеопотока в реальном времени.

1. Ronneberger, O. U-net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox. // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18. – Springer International Publishing, 2015. – P. 234–241.
2. Howard, A. Searching for MobileNetV3 / A. Howard [et al.] // Proceedings of the IEEE/CVF international conference on computer vision. – 2019. – P. 1314–1324.
3. Hu, J. Squeeze-and-excitation networks / J. Hu, L. Shen, G. Sun // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – P. 7132–7141.
4. Zhang, S. PortraitNet: Real-time portrait segmentation network for mobile device / S. Zhang [et al.] // Computers & Graphics. – 2019. – Vol. 80. – P. 104–113.
5. Accelerating TensorFlow Lite with XNNPACK Integration [Electronic resource] – Mode of access: <https://blog.tensorflow.org/2020/07/accelerating-tensorflow-lite-xnnpack-integration.html>. – Date of access: 16.10.2023.



Рис. 1 – Пример замены фона на видео. Слева направо: исходный кадр, результат сегментации, виртуальная замена фона

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ (СППР) КАК ИНСТРУМЕНТ ПРИНЯТИЯ ЭФФЕКТИВНЫХ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ НА ПРЕДПРИЯТИЯХ

Семижон Е. А.

Кафедра вычислительных методов и программирования, Белорусский государственный университет
информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: semijon.ekaterina@yandex.by

Информация играет важную роль в любой организации. Ценность информации зависит от ее применения и использования. Эффективность управленческих решений зависит главным образом от качества информации, используемой в качестве исходных данных в процессе принятия решений. Крайне важно, чтобы системы поддержки принятия решений своевременно и точно предоставляли информацию, помогающую менеджерам принимать правильные решения даже в непредвиденных ситуациях. В этой статье рассмотрены компоненты, виды систем поддержки принятия решений, а также ее использование как инструмента для принятия эффективных решений.

ВВЕДЕНИЕ

Эффективность деятельности предприятий зависит от качества управленческих решений. Современная практика изобилует многочисленными примерами ошибок в принятии управленческих решений со значительными отрицательными последствиями.

Необходимость выработки верных управленческих решений вызвана в общем виде рассогласованием между фактическим и требуемым или прогнозируемым состоянием развития объекта управления. Практика наглядно доказывает, что цена ошибки в сфере управления производством значительно выше, чем непосредственно в сфере самого производства.

Для облегчения процесса принятия управленческих решений была создана компьютерная автоматизированная система, целью которой является помощь людям, принимающим решение в сложных условиях для полного и объективного анализа предметной деятельности - **Система поддержки принятия решений (СППР)**.

I. ХАРАКТЕРИСТИКА И ВОЗМОЖНОСТИ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Система поддержки принятия решений – это автоматизированная система, которая поддерживает деловые или организационные действия по принятию решений. Системы поддержки принятия решений служат важным инструментом в управлении операциями и уровнями планирования, помогая менеджеру и сотрудникам предприятия принимать решения по сложным проблемам, которые могут быстро меняться и не выявляться заранее.

Системы поддержки принятия решений могут быть автоматизированными или автоматическими.

Система поддержки принятия решений обладает следующими основными характеристиками:

- Обработка больших объемов данных;

- Получение и обработка данных из различных источников, включая внутренние и внешние данные, хранящиеся в системах и сетях мэйнфреймов;
- Обеспечение гибкости отчетов и презентаций в соответствии с потребностями лица, принимающего решения;
- Повышение эффективности принимаемых решений;
- Наличие прогностических инструментов.

Современные виды систем поддержки принятия решений наделены следующими возможностями:

- Поддержка этапов решения проблем, включая анализ, проектирование, выбор, внедрение и мониторинг;
- Поддержка различных частот принятия решений, от единичных до повторяющихся;
- Поддержка различных типов проблем, от высокоструктурированных до неструктурированных;
- Поддержка различных уровней принятия решений, включая решения оперативного, тактического, стратегического уровня.

II. КОМПОНЕНТЫ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Система поддержки принятия решений требует трех первичных компонентов: модели управления; управления данными для сбора и ручной обработки данных; пользовательский интерфейс. Они показаны на рисунке 1.

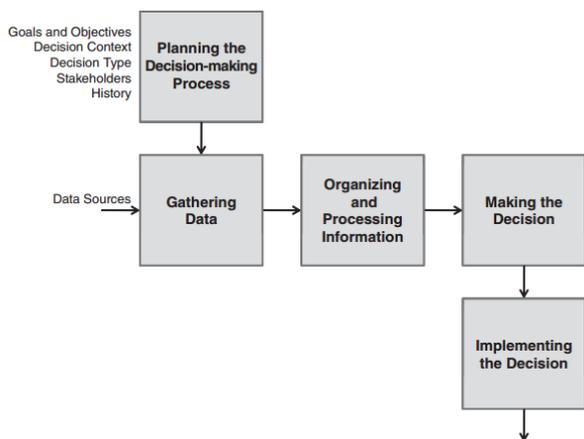


Рис. 1 – Компоненты системы поддержки принятия решений.

III. ВИДЫ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Различают 5 видов компьютерных СППР:

- Коммуникативные. Ориентированы на одновременную работу нескольких специалистов, занятых одной общей задачей.
- Информационные. Сосредоточены на сборе и обработке данных, преимущественно анализе временных рядов, функционируют, как СУБД в пределах одной компании.
- Документальные. Предназначены для обработки и анализа документов разного формата со структурированными и не структурированными данными.
- Интеллектуальные. Содержат данные о решениях аналогичных задач, нормы и правила, на основании которых они принимались, предлагают готовые алгоритмы, исходя из накопленного опыта.
- Моделируемые. Подбирают модели бизнес-процессов по заданным условиям – статистические, финансовые, аналитические).

IV. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ ИСПОЛЬЗОВАНИЯ СИСТЕМ ПРИНЯТИЯ РЕШЕНИЯ

Использование систем поддержки принятия решений на предприятиях с одной стороны имеет много преимуществ, а с другой стороны имеет несколько недостатков, которые заключаются в следующем:

Преимущества:

1. Повышает организационный контроль за счет постоянного предоставления менеджерам информации, которая может быть ис-

пользована для оценки производительности сотрудников;

2. Повышает эффективность и скорость принятия решений; - помогает автоматизировать управленческие процессы;
3. Повышает способность организации к прогнозированию и, следовательно, способность принимать решения на будущее;
4. Улучшает межличностную коммуникацию внутри организации.

Недостатки:

1. Стоимость разработки и внедрения СППР требует огромных капиталовложений, что делает ее менее доступной для небольших организаций;
2. Менеджеры могут слишком полагаться на систему, что устраняет субъективный аспект принятия решений;
3. В некоторых случаях пользователи могут получить избыточную информацию о проблеме, что может снизить эффективность принятого решения.

Несмотря на недостатки систем поддержки принятия решений, их существование как одного из элементов интеллектуального капитала на предприятии, и их использование в процессе принятия управленческих решений стало одним из необходимых факторов для продолжения деятельности предприятия, и ее выживания в конкурентной борьбе, особенно в свете промышленного развития и большого технического прогресса, а также сложности окружающей среды предприятия и огромного количества внутренних и внешних данных. Где системы поддержки принятия решений повышают качество и эффективность принимаемых решений и облегчают работу менеджеров разного уровня, а также увеличивают скорость решения проблем в бизнесе.

СПИСОК ЛИТЕРАТУРЫ

1. Alyoubi B.A. Decision support system and knowledge-based strategic management // ELSEVIER. 2015. № 65 – С 278-284.
2. Висайтова Л.Р. Системы поддержки принятия решений: эффективность принятия решений // Научно-практический электронный журнал Аллея Науки. 2020. №11(50) – С 219- 225.
3. Tripathi K.P. Decision support system is a tool for making better decisions in the organization // Indian Journal of Computer Science and Engineering. 2017. № 21– С 112-117.
4. Khodashahri N.G, Sarabi M.H. Decision support system (DSS) // Singaporean journal of business economics and management studies. 2013. №6 – С 94-102

НЕЙРОННЫЕ СЕТИ ДЛЯ ПРОГНОЗИРОВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ

Скиба И. Г., Нестеренков С. Н., Внук О. М.

Отдел информационных технологий,

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

Минск, Республика Беларусь

E-mail: {i.skiba, s.nesterenkov, o.vnuk}@bsuir.by

Прогнозирование результатов учебного процесса в высшем учебном заведении является важной задачей в образовании. В данной статье было рассмотрено применение нейронных сетей для прогнозирования результатов учебного процесса в высшем учебном заведении. В данном исследовании были рассмотрены различные подходы к созданию нейронных сетей и оценена их эффективность в прогнозировании результатов учебного процесса. Были выявлены преимущества и недостатки каждого подхода и даны рекомендации по выбору наиболее подходящего для конкретных задач. В итоге, были получены прогнозы, которые позволяют определить студентов, испытывающих трудности в учебе. Нейронная сеть проявила высокую точность в прогнозировании результатов учебного процесса, что дает надежду на более эффективное управление образовательным процессом в высших учебных заведениях.

ВВЕДЕНИЕ

Прогнозирование результатов учебного процесса является одной из наиболее актуальных задач, с которыми сталкиваются высшие учебные заведения по всему миру. Предсказание успеха студентов может помочь университетам определить тех студентов, которые нуждаются в дополнительной поддержке, и предоставить им индивидуальный подход к обучению.

Нейронные сети являются мощным инструментом для анализа и обработки данных, что позволяет прогнозировать результаты учебного процесса на основе различных факторов, таких как история успеха студента, методы обучения и даже личные характеристики студента. [1]

Исследование посвящено возможности использования данных о студентах для прогнозирования их успеха в учебе. В данной статье рассмотрены различные источники данных, такие как оценки, учебные планы, история успеха и личные характеристики студентов, и будет оценена их значимость для прогнозирования результатов обучения.

В итоге, данное исследование будет способствовать повышению качества образования в высших учебных заведениях, улучшению процесса прогнозирования результатов обучения и предоставлению студентам более эффективной поддержки в обучении.

I. ВЫДЕЛЕНИЕ ПРИЗНАКОВ

Выделение признаков – это процедура отбрасывания незначимых переменных из очищенной выборки перед запуском машинного обучения и интеллектуального анализа данных [2]. Сокращение числа предикторов необходимо по нескольким причинам:

- Значимость признаков – как правило, исходная выборка всегда содержит много «му-

сорных данных»: шумов, выбросов, а на реальный результат влияют лишь несколько предикторов [3].

- Точность решения – некоторые модели машинного обучения чувствительны к величине входного вектора. Например, у нейронных сетей большое число входных данных может привести к переобучению.
- Скорость вычислений – чем меньше переменных, тем быстрее будут идти расчеты.

Таким образом, снижение размерности задачи – необходимый этап подготовки данных, который оказывает решающее значение на итоговый результат.

Среди полученных данных находится большое количество информации о каждом студенте. Задача выявления признаков заключается в нахождении данных, влияющих на успешное завершение обучения.

II. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Прогнозирование вероятности успешного завершения обучения студентов является важной задачей в образовании, которая может быть решена с помощью различных методов машинного обучения. В данной статье были рассмотрены три модели нейронной сети: многослойный персептрон, модель сверточной нейронной сети и модель рекуррентной нейронной сети. Ниже будут приведены их преимущества и недостатки в контексте прогнозирования вероятности успешного завершения обучения студентов по разным признакам.

Многослойный персептрон – это одна из самых популярных моделей нейронных сетей, которая используется для решения широкого круга задач, включая классификацию, регрессию и прогнозирование. В этой модели нейроны организованы в несколько слоев, каждый из которых

имеет свои входы и выходы [4]. Основным преимуществом данной модели является способность к обучению на больших объемах данных с высокой точностью [5].

Сверточная нейронная сеть – это тип нейронной сети, который используется для анализа и обработки изображений, видео, звука и других типов сигналов. Она была разработана для эффективной обработки больших объемов данных и получила широкое применение в таких областях, как компьютерное зрение, распознавание речи, обработка естественного языка и многих других [6].

Рекуррентная нейронная сеть – это тип нейронной сети, который используется для анализа последовательностей данных, таких как текст, звук, временные ряды и другие. Она была разработана для обработки данных, которые имеют зависимости во времени, что делает ее особенно полезной в таких областях, как распознавание речи, машинный перевод, генерация текста и многих других [7].

Из всех трех типов нейронных сетей, многослойный перцептрон является наиболее подходящим для решения задачи прогнозирования вероятности успешного завершения обучения на основе числовых признаков, таких как средний балл студента, результаты сдачи ЦТ, количество пропусков и пересдач. Многослойный перцептрон хорошо работает с данными, которые имеют множество числовых признаков.

III. ЭКСПЕРИМЕНТЫ

За входные данные для обучения нейронной сети была взята база данных студентов за 2 года. Для обучения было взято около тысячи студентов, закончивших обучение. Для студентов, которые выпустились, выходные данные принимались за 1, для студентов, которые были отчислены, - 0. После обучения нейронной сети, через нее было проведено несколько студентов, которые еще обучаются. В результате была получена вероятность успешного завершения обучения.

Результаты предсказаний нейронной сети представлены в таблице 1.

Таблица 1 – Результаты эксперимента

Средний балл	Баллы ЦТ	Количество пересдач	Часы пропусков	Результат
8.36	247	2	525	28%
8.06	251	0	23	78%
8.41	248	1	12	75%
9.09	253	1	31	75%
5.67	247	0	0	78%
6.44	252	0	176	65%

По результатам, приведенным в таблице, можно заметить, что большое влияние на результат имеет количество пропусков. Например, у студента №1 525 академических часов пропуска, что

объясняет такую низкую вероятность успешного завершения обучения. Также можно отметить, что у студента №3 небольшое количество пропусков, но при этом результат не максимальный.

IV. ЗАКЛЮЧЕНИЕ

Нейронные сети – это мощный инструмент для прогнозирования результатов обучения, а также анализа больших объемов данных для выявления скрытых закономерностей и прогнозирования будущих результатов. Они могут использоваться в различных областях, где необходимо принимать решения на основе исторических данных, таких как образование, финансы, здравоохранение и многие другие.

Однако, создание эффективной нейронной сети требует правильного выбора архитектуры, алгоритма обучения и гиперпараметров. Кроме того, необходимо иметь достаточное количество данных для обучения и тестирования нейронной сети, чтобы получить точные результаты. Важно помнить, что нейронные сети не являются универсальным решением для всех проблем и задач, и их эффективность зависит от качества данных и правильного подхода к обучению.

Кроме того, выбор гиперпараметров, таких как скорость обучения, количество слоев и количество нейронов, также имеет большое значение для эффективности работы нейронной сети.

СПИСОК ЛИТЕРАТУРЫ

1. Скиба, И. Г. Анализ образовательных данных в высших учебных заведениях / И. Г. Скиба, И. С. Тарасюк, С. Н. Нестеренков // Информационные технологии и системы 2022 (ИТС 2022). – Минск, 2022. – С. 167-168.
2. Мартышенко, С. Н. Автоматизация анализа данных в исследовании социально-экономических процессов / С. Н. Мартышенко, Л. С. Мазелис, К. С. Солодунин. – Владивосток: Изд-во ВГУЭ, 2019.. – С. 164.
3. Шмыгарева, В. С. Разработка нейронных сетей / В. С. Шмыгарева // Молодой ученый. – 2020. – №. 24 (419). – С. 79-87.
4. Подзорова М. И. Нейронная сеть как одно из перспективных направлений искусственного интеллекта / М. И. Подзорова // Modern European Researches. – 2022. – Т. 1. – №. 3 – С. 169-176.
5. Фостер, Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. / Д. Фостер. – Санкт-Петербург: Питер, 2020.. – С. 336.
6. 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) = Information Technologies and Systems 2017 (ITS 2017) : материалы междунар. науч. конф. (Республика Беларусь, Минск, 25 октября 2017 года) – Минск : БГУИР, 2017. – С. 212.
7. Нестеренков, С. Н. Применение больших данных в электронном образовании. / С. Н. Нестеренков, М. И. Макаров, Н. В. Ющенко, А. Д. Радкевич // Пятая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», – Минск, 2019.

ПРИМЕНЕНИЕ МЕТОДА СИСТЕМНОГО АНАЛИЗА ПРИ ПРОЕКТИРОВАНИИ АВТОМАТИЗИРОВАННЫХ СИСТЕМ

Струкова А. А., магистрант, Шилин Л. Ю., Хмыз Д. Д.
Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: alinastrukovaa@gmail.com, dekfitu@bsuir.by

Разработана автоматизированная информационная система, использующая метод системного анализа для выбора наиболее подходящих альтернатив в условиях многокритериальности. Разработаны алгоритмы подбора альтернатив для пользователей на основе данных их профилей и предоставленной информации. Реализован подбор рекомендаций в системе для разных типов пользователей.

ВВЕДЕНИЕ

Автоматизированные информационные системы являются одной из важнейших тем в современном управлении ресурсами человеческого потенциала. В условиях быстрого развития технологий и необходимости внедрения новых методов в системы, эффективное использование автоматизированных систем становится ключевым фактором успеха организаций. Внедрение автоматизированной системы с использованием метода системного анализа позволит экономить временные, денежные и человеческие ресурсы на поиске и подборе необходимых данных, автоматизировать процессы.

I. ПРИМЕНЕНИЕ МЕТОДА СИСТЕМНОГО АНАЛИЗА

Математическое обеспечение в разрабатываемой автоматизированной информационной системе представлено модифицированным алгоритмом Кемени-Снелла. Данный алгоритм относится к методам системного анализа и основан на принципах транзитивности и согласованности в предпочтениях.

Модифицированный алгоритм Кемени-Снелла предназначен для ранжирования альтернатив, оцениваемых по критериям различных видов: числовым, качественным, «да» – «нет», с учетом суждений эксперта о важности критериев. Алгоритм основан на ранжировании и попарном сравнении альтернатив по каждому критерию. Задачи, решаемые данным алгоритмом, можно отнести к задачам принятия решения при многих критериях. В данном случае задачи являются слабоструктурированными, так как в их постановке имеются как объективные данные, так и субъективные.

Рассмотрим применение алгоритма Кемени-Снелла для автоматизированной системы подбора персонала, что позволит подбирать для соискателей подходящие вакансии на основе их резюме и данных, а для работодателей подходящих кандидатов на основании указанной в их вакансиях информации.

Существует задача о подборе подходящих вакансий для соискателя с помощью рассматриваемого алгоритма. Предположим, есть список альтернатив по трём вакансиям. Альтернативами в данном случае будут являться соответствия между профилем соискателя и вакансиями, содержащими разную информацию. Данные для подбора подходящих вакансий для соискателя представлены в таблице 1.

Здесь С1, С2, С3 - вакансии. Таблица отображает соответствие между данными соискателя и информацией по каждой вакансии.

Таблица 1 – Данные для подбора подходящих вакансий для соискателя

Критерий	С1	С2	С3
Совпадение по названию вакансии, %	55	33	43
Совпадение по длительности рабочего дня	Не совпадает	Совпадает	Совпадает
Суммарное совпадение по названиям технологий	Выше, чем указано в вакансии	Ниже, чем указано в вакансии	Совпадает

Воспользуемся методом Саати для нахождения весов критериев, представляющих собой числовые оценки их важности. Обозначим критерии:

- К1 – совпадение по названию вакансии;
- К2 – совпадение по длительности рабочего дня;
- К3 – суммарное количество совпадений по названиям технологий;
- К4 – совпадение по зарплате.

Попарное сравнение критериев для данной задачи представлено в таблице 2.

Таблица 2 – Оценка критериев по методу Саати

	К1	К2	К3	К4
К1	1	5	2	3
К2	1/5	1	1/3	1
К3	1/2	3	1	3
К4	1/3	1	1/3	1

Далее выполним попарное сравнение критериев и, выполнив обработку матрицы парных сравнений, найдём веса критериев: $V_1 = 0,47$; $V_2 = 0,1$; $V_3 = 0,3$; $V_4 = 0,12$. Выполняется ранжирование альтернатив по каждому из критериев.

На основе ранжирования альтернатив по каждому из критериев составляется матрица парных сравнений. Правила заполнения матриц парных сравнений в модифицированном алгоритме Кемени-Снелла описывают, какие значения присваиваются лучшим и худшим альтернативам. В результате анализа матриц парных сравнений каждая альтернатива сравнивается со всеми другими.

Составляется матрица потерь. Элементы матрицы потерь считаются по следующей формуле:

$$R_{jk} = \sum_{i=1}^M V_i \cdot |R_{jk}^i - 1|, \quad j = 1, \dots, N, \quad k = 1, \dots, N.$$

Матрица потерь для данной задачи представлена в таблице 3.

Таблица 3 – Матрица потерь

	C1	C2	C3
C1	-	0,44	0,44
C2	1,54	-	1,58
C3	1,54	0,4	-

Размерность матрицы – $N \times N$, где N – количество альтернатив. Смысл элементов матрицы потерь следующий: чем больше элемент R_{jk} , тем больше отставание j -й альтернативы от k -й (тем хуже j -я альтернатива по сравнению с k -й).

Выполняется предварительное ранжирование альтернатив. Для этого находятся суммы строк матрицы потерь. Смысл этих сумм следующий: сумма j -й строки представляет собой оценку отставания j -й альтернативы от всех остальных альтернатив.

Альтернатива, которой соответствует минимальная сумма, предварительно считается лучшей. Строка и столбец этой альтернативы исключаются из матрицы потерь. Суммирование строк матрицы потерь и исключение альтернатив выполняются до тех пор, пока не будет исключена вся матрица. Чем раньше исключена альтернатива, тем она лучше.

Выполним предварительное ранжирование. Найдём суммы строк матрицы потерь:

$$P_1 = 0,44 + 0,44 = 0,88;$$

$$P_2 = 1,54 + 1,58 = 3,12;$$

$$P_3 = 1,54 + 0,4 = 1,94.$$

Предварительно считается лучшей альтернатива C1. Она исключается из матрицы потерь. Составляется сокращённая матрица потерь, суммы строк которой:

$$P_2 = 1,58;$$

$$P_3 = 0,4;$$

Лучшая альтернатива из двух оставшихся – C3.

Предварительное ранжирование альтернатив: C1, C3, C2.

Затем выполняется окончательное ранжирование альтернатив. Для этого альтернативы сравниваются попарно, начиная с конца предварительного ранжирования.

Сравниваем C3 и C2. $R_{32}=0,4$; $R_{23}=1,58$. Так как $R_{32} < R_{23}$, альтернативы остаются на своих местах.

Сравниваем C1 и C3. $R_{13}=0,44$; $R_{31}=1,54$. Так как $R_{13} < R_{31}$, альтернативы остаются на своих местах.

Окончательное ранжирование альтернатив для решаемой задачи следующее: C1, C3, C2.

На основе рассмотренного примера можно сделать вывод о необходимости внедрения алгоритма Кемени-Снелла в автоматизированные системы, так как это позволит наиболее точно подбирать подходящие альтернативы. Рассматриваемый метод системного анализа может быть использован в автоматизированных информационных системах для выполнения следующих задач:

- принятие решений;
- ранжирование;
- оценка производительности;
- подбор рекомендаций и альтернатив.

II. ЗАКЛЮЧЕНИЕ

Исследования в области автоматизированных информационных систем с использованием методов системного анализа имеют большое значение для разных сфер. Актуальные тенденции и перспективы развития систем требуют дальнейших исследований и улучшений, чтобы максимально использовать потенциал автоматизации в подборе наиболее подходящих альтернатив в условиях многокритериальности и достичь наилучших результатов для организаций.

СПИСОК ЛИТЕРАТУРЫ

1. Набатова, Д. С. Математические и инструментальные методы поддержки принятия решений : учебник и практикум для вузов / Д. С. Набатова. – Москва : Издательство Юрайт, 2020. – 292 с.
2. Назаров, С. В. Архитектуры и проектирование программных систем : монография / С. В. Назаров. – М.: ИНФРА-М, 2013. – 413 с.

МЕТОДЫ ВОССТАНОВЛЕНИЯ НЕПОЛНЫХ ДАННЫХ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ АНАЛИЗА

Судаков Б. Д., Герман О. В.

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

E-mail: sudakov.bogdan666@gmail.com, ovgerman@bsuir.by

В данной работе рассматриваются методы восстановления неполных данных, которые позволяют повысить точность анализа и улучшить качество получаемых результатов. Представлен обзор различных подходов, включая регрессионный анализ, метод хот-дек и алгоритм ZET. Обсуждается применимость каждого метода в зависимости от специфики задачи и характера неполноты данных.

ВВЕДЕНИЕ

При проведении статистического анализа на практике ограничиваются анализом не всей генеральной совокупности в целом, а лишь некоторого выборочного числа наблюдений. Анализируемая выборка должна отвечать критериям качества и полноты. В реальности приходится сталкиваться с ситуацией, когда некоторые из свойств одного или нескольких объектов отсутствуют – возникает ситуация данных с пропусками, что значительно осложняет математическую обработку, так как смещение основных статистических характеристик, таких как математическое ожидание или дисперсия, например, возрастает прямо пропорционально числу пропусков. На сегодняшний день в математической статистике существует несколько путей решения проблемы неполных данных:

- Исключение некомплектных объектов из исходной выборки;
- Применение специально разработанных математических методов анализа неполных данных, таких как метод взвешивания или метод максимального правдоподобия и EM-алгоритм [1];
- Восстановление пропусков (наиболее распространены методы заполнения по среднему и по регрессии).

1. МЕТОДЫ ВОССТАНОВЛЕНИЯ ПРОПУСКОВ

Согласно классификации Литтла и Рубина [2], в данных могут встречаться:

1. Полностью случайные пропуски (missing completely at random (MCAR));
2. Случайные пропуски (missing at random (MAR));
3. Неигнорируемые пропуски (non-ignorable missingness).

С полностью случайными пропусками (MCAR) мы имеем дело тогда, когда пропуски случайно распределены в массиве данных по всем переменным. Наличие MCAR можно проверить статистически t-тестом или хи-квадрат тестом. В модуле SPSS Missing Values Analysis (MVA) есть опция Little's MCAR test, которая на базе статисти-

стики хи-квадрат проверяет данные на MCAR. Если в тесте наблюдается незначимый уровень критерия, то мы имеем дело с полностью случайными пропусками данных. Случайные пропуски (MAR) данных встречаются тогда, когда пропуски в массиве данных случайно распределены не по всем переменным, а только внутри каких-либо определенных подгрупп переменных. Такое распределение пропусков в данных случается гораздо чаще, чем MCAR.

Для борьбы с этими двумя видами пропусков применяют восемь основных классов методов:

1. Анализ полных наблюдений (listwise deletion);
2. Методы, использующие доступную информацию (pairwise deletion);
3. Подстановка среднего по выборке (mean substitution);
4. Метод хот-дек (hot deck);
5. Регрессионный анализ (regression);
6. Оценка с помощью максимизации правдоподобия (maximum likelihood estimation);
7. Подстановка с помощью факторного анализа (factor analysis substitution);
8. Модель множественного восстановления данных (multiple imputations method).

Два первых метода широко распространены в исследовательской практике.

Методы с 3-го по 7-ой используют принцип однократной подстановки восстановленных тем или иным способом данных и могут использоваться, если пропуски распределены случайно (MAR). В основе методов лежит принцип вычисления и подстановки взамен каждого пропущенного значения одного нового значения. Кратко рассмотрим эти методы.

Метод хот-дек (hot deck) представляет собой метод подстановки среднего по выборке с некоторым количеством модификаций. Наиболее простой вариант — сортировка респондентов по ключевым переменным, тогда респонденты со схожими ответами находятся рядом друг с другом [3]. В качестве ключевых чаще всего выступают социально-демографические переменные, но можно использовать и другие переменные, имеющие корреляции с переменной с пропущенными

данными. При восстановлении пропущенные значения переменной заимствуются из предыдущего наблюдения.

Регрессионный анализ. В зависимости от типа данных используются либо множественная линейная, либо логистическая регрессия. На базе наблюдений, не содержащих пропущенных данных, вычисляются коэффициенты регрессии, и далее с их помощью восстанавливается пропущенное значение зависимой переменной. Если не говорить об обычных проблемах регрессии, таких как мультиколлинеарность, гомоскедастичность и т.д., то можно обозначить две проблемы, связанные с её использованием для восстановления пропущенных данных. Во-первых, из-за самой природы регрессии мы полностью исключаем случайные вариации. Это ведет к тому, что при большой доле пропущенных значений становится очень заметным смещение результатов по направлению к средним оценкам. Для борьбы с этим используется метод случайной подстановки, при котором к вычисленному значению прибавляются случайные величины. Во-вторых, используя в уравнение регрессии слишком большой набор независимых переменных, мы рискуем моделировать шум вместо каких-то осмысленных значений переменных.

Алгоритм ZET является детально проработанной и апробированной технологией верификации экспериментальных данных, основанной на гипотезе их избыточности. Главная идея алгоритма ZET заключается в подборе «компетентной матрицы», используя данные из нее находят параметры зависимости, которая применяется для прогнозирования пропущенного значения. Субъективизм определения размерности «компетентной матрицы» приводит к учету неинформативных и шумовых факторов и смещению оценки неизвестного значения. Основное отличие алгоритма состоит в определении оптимального размера «компетентной матрицы». Рассмотрим подробнее методику алгоритма ZET.

II. АЛГОРИТМ ZET

В основе алгоритма ZET лежат три предположения. Первое (гипотеза избыточности) состоит в том, что реальные таблицы имеют избыточность, проявляющуюся в наличии похожих между собой объектов (строк) и зависящих друг от друга свойств (столбцов). Если же избыточность отсутствует (как, например, в таблице случайных чисел), то предпочесть один прогноз другому невозможно. Второе предположение (гипотеза локальной компактности) состоит в утверждении, что для предсказания пропущенного элемента нужно использовать не всю таблицу, а лишь ее «компетентную» часть, состоящую из элементов

строк, похожих на строку i , и элементов столбцов, похожих на столбец j . Остальные строки и столбцы для данного элемента неинформативны. Их использование лишь разрушало бы локальную компактность подмножества компетентных элементов и ухудшало точность предсказания. Третье предположение (гипотеза линейных зависимостей) заключается в том, что из всех возможных видов зависимостей между столбцами (строками) в алгоритме ZET используются только линейные зависимости. Если зависимости носят более сложный характер, то для их надежного обнаружения требуется такой большой объем данных, который в реальных задачах встречается нечасто. Для различных прикладных задач были сделаны многочисленные модификации базового алгоритма ZET, отличающиеся своим назначением и наборами разных режимов работы. Программы заполнения пробелов могут работать в одном из следующих режимов:

- Заполнение всех пробелов;
- Заполнение только тех пробелов, ожидаемая ошибка для которых не превышает заданной величины;
- Заполнение пробелов только на базе информации, имеющейся в исходной таблице;
- Заполнение каждого следующего пробела с использованием исходной информации и прогнозных значений ранее заполненных пробелов.

III. ВЫВОДЫ

В данной работе рассмотрены методы восстановления неполных данных, которые позволяют повысить точность анализа и улучшить качество принимаемых решений. Применение алгоритмов восстановления неполных данных позволяет избежать потери важной информации, которая может негативно повлиять на точность анализа. Кроме того, использование этих алгоритмов помогает снизить риск ошибок, связанных с пропуском важных данных. Таким образом, алгоритмы восстановления неполных данных являются важным инструментом для повышения точности и надежности анализа данных, а также обеспечения более качественного принятия решений на основе полученных результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Rubin, D. B. *Statistical Analysis with Missing Data* / D. B. Rubin, R. J. A. Little // New York: John Wiley & Sons – 2017. – P. 40–43.
2. Wedel, M. *Factor Analysis and Missing Data* / M. Wedel, W. A. Kamkura // *Journal of Marketing Research*. – 2000. – № 4. – С. 490–498.
3. Косьяненко, А. В. Опыт восстановления пропущенной рыночной информации на основе Байесовского подхода / А. В. Косьяненко // *Издательство: Инфра-М*, 2016. – 193 с.

ВОПРОСЫ ИНТЕГРАЦИИ СРЕДСТВ LMS MOODLE И СИСТЕМЫ WOLFRAM MATHEMATICA ПРИ СОЗДАНИИ И РАЗВИТИИ ИНТЕЛЛЕКТУАЛЬНОГО КОНТЕНТА

Таранчук В. Б.

Кафедра компьютерных технологий и систем, Белорусский государственный университет

Минск, Республика Беларусь

E-mail: taranchuk@bsu.by

Обсуждаются методические и технические решения, варианты интеграции средств LMS Moodle и системы Wolfram Mathematica при создании и развитии адаптивного интеллектуального контента, применяемого при преподавании дисциплин компьютерного моделирования, интеллектуального анализа данных.

ВВЕДЕНИЕ

В настоящее время рамках актуальных образовательных проектов ведущих университетов реализуются программы активного использования интеллектуальных образовательных технологий (ИОТ), основой которых является создание в процессе учебно-исследовательской деятельности новых интеллектуальных продуктов, обеспечивающих поддержку учебного процесса, неразрывные связи решаемых учебных, научных и практических задач, многоплановость сотрудничества преподавателей и студентов. Формирование у обучаемых различных новых приемов генерирования и воспроизводства современных знаний обеспечиваются инструментами ИОТ. При этом спектр возможностей очень широк – средствами интеллектуальных технологий реализуются, в частности, парадигмы образования на протяжении всей жизни (lifelong learning), мобильного обучения (m learning), обучения, проникающего во все сферы жизни общества и человека (u learning, ubiquitous learning), открытого и смешанного (очно-виртуального) образования.

Использование электронных образовательных ресурсов стало нормой практически на всех этапах и для всех видов обучения. Программное обеспечение ИОТ предоставляет эффективные инструменты создания и сопровождения контента, нет проблем с дружественным интерфейсом, но в отдельных случаях этого уже недостаточно. Нужны инструменты интеллектуальной адаптации, которые позволили бы обучаемым, дали желаемую возможность маневра, свободного выбора собственной траектории обучения, причем (что принципиально важно), при минимальном вмешательстве преподавателя.

В докладе и материалах конференции планируется обсудить новые апробированные методические и технические решения, приёмы адаптации компьютерных средств, методы и инструменты создания и сопровождения интерактивных интеллектуальных образовательных ресурсов. Примерами иллюстрируются варианты применения предлагаемого подхода создания умной образовательной среды, при очно-виртуальном преподавании ряда дисциплин в БГУ на факультете

прикладной математики и информатики. Конкретно, знаковые позиции технологии обозначены при преподавании дисциплины «Компьютерный анализ и визуализация» [1]. Базовые позиции, вопросы развития и адаптации интерактивной технологии обучения типа eLearning в среде Moodle; варианты расширения базовых функциональных возможностей Moodle путем включения дополнительных сервисов, интерактивных ресурсов формата вычисляемых документов CDF, как они технически реализованы было пояснено в [2].

I. О ФОРМАТЕ ЗАНЯТИЙ

Несколько существенных аспектов проведения занятий, применения, настройки, адаптации интеллектуальных средств обучения.

Лекции и лабораторные занятия проводятся один раз в неделю в компьютерном классе, и обязательно – согласно расписанию лабораторные занятия проходят сразу за лекционными. Все занятия поддерживаются системой электронного обучения (СЭО) Moodle. В каждом конкретном ресурсе СЭО «Курс» в первом блоке «Общее» размещены элементы «Объявления», форум «Вопросы и ответы по заданиям лекций, тестов», страница «О дисциплине». Далее размещаются блоки «К лекции, практике N (дд/мм)», причем, с сортировкой «Новый сверху». Содержание таких блоков поясняется отдельно.

Фактически все студенты работают на личных компьютерах, доступ к Сети – университетский или от своих провайдеров (предпочтения у большинства именно такие). В течение всего курса действует рейтинговая система, каждый студент имеет индивидуальный рейтинг, таблица рейтингов доступна всем. Соответствующие баллы рейтингов обновляются не реже одного раза в неделю, публикуются преподавателем на странице элемента «Форум» Moodle. Баллы включают призовые за правильные ответы на блиц-вопросы на лекциях, штрафные баллы за пропуски лекций (в случае уважительных причин такие обнуляются), баллы по итогам тестов, призовые баллы за инициативные предложения уточнений материалов лекций и тестовых вопросов/заданий.

Регламент всех циклов занятий одинаков. Материалы предстоящей лекции, расширенный конспект в PDF документе, ссылки на рекомендуемые материалы по теме выкладываются на страницах элемента «Лекция» Moodle за день до занятия. Также на таких страницах записываются коды программных модулей упражнений и примеров демонстрируемых «живьём» на лекциях (ASCII формат, textual representations для nb исходных кодов Wolfram *Mathematica*).

На каждом лекционном занятии большинство времени – показ «живьём» работы программных модулей, кодов программ по изучаемым темам, часто с демонстрацией типовых ошибок и пояснениями «что не по правилам»; также – акценты по ключевым абзацам конспектов. Работающие в аудитории студенты (большинство) повторяют на своих компьютерах по ходу лекции показываемое «живьём» преподавателем (работает на своем компьютере, и его экран дублируется на панели или проектором). Все в аудитории в ходе лекции также имеют возможность получить призовые баллы за правильные ответы на вопросы/задания, которые «всплывают» в разные моменты. Таковые в модуле занятия в блоке «К лекции, практике» скрыты на старте, открываются в предусмотренное время. Как правило, это короткие вопросы по темам рассказанного или короткие упражнения, предполагающие вычисления по изучаемым программным кодам. Вопросы/задания формулируются в элементах «Тест» Moodle, сопровождаются паролем (озвучивает преподаватель после «всплывают»). Такие вопросы доступны 2-3 минуты, дистанционные ответы не возможны, т.к. надо знать пароль. Часть вопросов «2-3 минуты» предполагает – найти альтернативные формулировки, интерпретации в Сети. По всем вопросам/заданиям в выдаваемой элементом «Тест» «Категория» не менее 7 вариантов, вопросы от СЭО идут по правилу «Случайный», поэтому в большинстве случаев сидящие рядом получают разные вопросы. Никаких ограничений доступа студентов в поисковые системы, обмена ответами в социальных сетях со стороны преподавателя нет. На каждой лекции выкладываются 2-3 вопроса/задания, за каждый правильный ответ (с учетом сложности) рейтинг пополняется на 3-7 баллов, причем, такие баллы призовые, неправильный ответ – 0 баллов.

Отметим основные выводы из оценок итогов практики преподавания последних лет по вопросам объема материала и качества выкладываемых в СЭО конспектов на страницах «Лекция». По ходу лекции или после занятий студенты на форуме размещают вопросы по изучаемому материалу – на них ответы размещаются там же, или

более подробно обсуждаются на следующей лекции. Статистика – примерно по 1/4 тем студенты запрашивают консультации или дополнительные пояснения. Студенты авторы интересных вопросов (по непонятно изложенным материалам или не включенным в конспект) получают призовые баллы в рейтинг. Важно мотивировать такую взаимосвязь, чтобы учитывать мнения обучаемых при подготовке следующего конспекта, тестовых заданий.

II. КОМПЬЮТЕРНЫЕ ТЕСТЫ

Курс включает 16 обязательных тестов, соответственно, контрольный уровень по всей дисциплине – 1600 баллов. Если по дисциплине предусмотрен зачет, допуск автоматически получают все, у кого личный рейтинг составляет не менее 2/3 от 1600. Итоги каждого теста анализируются с использованием инструментов СЭО «Рассчитать статистику», с особым вниманием к позиции «Индекс легкости». Именно это обеспечивает адаптацию образовательного маршрута в каждой конкретной группе обучаемых. Основное, как это делается? Каждый тест структурно содержит 3 части: вопросы по терминологии предметной области в соответствии с образовательным стандартом специальности (компетентность) – по баллам не более 25; задания прошлого теста, которые по показателю «Индекс легкости» оказались самыми не успешными – по баллам около 30; задания (не терминология или определения) по изучаемым темам текущего занятия – примерно на 30-40 баллов. Дополнительная особенность – каждый тест контролируется по показателю среднего балла участников, когда в расчете среднего не учитываются один максимальный и один минимальный баллы. Такие тесты (классифицируются – «не обеспечены качественными пояснениями на лекции») предлагаются для повторения желающим в первые 40 минут следующего лабораторного занятия. Из практики 2022 года – из 16 тестов повторили только 2.

СПИСОК ЛИТЕРАТУРЫ

1. Компьютерный анализ и визуализация: учебная программа учреждения высшего образования по учебной дисциплине для специальности: 1-31 03 07 Прикладная информатика. / В.Б. Таранчук // [Electronic resource]. – Mode of access: <https://www.elib.bsu.by/handle/123456789/290857>. – Date of access: 16.10.2023.
2. Таранчук, В. Б. Практические аспекты разработки, сопровождения, использования интеллектуальных информационно-образовательных ресурсов / В. Б. Таранчук // Информатизация образования и методика электронного обучения : материалы III Международ. науч. конф. Красноярск, 24–27 сентября 2019 г. : в 2 ч. Ч. 1 / Красноярск : Сиб. федер. ун-т, – 2019. – С. 116–121.

ИСПОЛЬЗОВАНИЕ ГРАДИЕНТНОГО БУСТИНГА НАД РЕШАЮЩИМИ ДЕРЕВЬЯМИ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Хмельницкий В. В.

Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: vovchyt@gmail.com

В данной статье рассматривается применение градиентного бустинга для решения задачи прогнозирования временных рядов. Приводится краткое описание модели машинного обучения, а также сравнение со статистическими моделями, решающими аналогичную задачу.

ВВЕДЕНИЕ

Временные ряды являются важной частью современной аналитики данных и имеют большое практическое применение в различных областях, начиная от финансов и заканчивая метеорологией.

Существует множество способов прогнозирования временных рядов, однако многие из них имеют свои ограничения, среди которых большие затраты ресурсов и времени, а также необходимость в дополнительной обработке входных данных. Применение градиентного бустинга может помочь обойти часть данных ограничений.

I. ОСНОВНЫЕ ПРОБЛЕМЫ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

На данный момент существует большое количество моделей, прогнозирующих временные ряды, но они делятся на две основные категории: статистические и машинное обучение. Особой популярностью среди статистических моделей пользуется ARIMA (Autoregressive Integrated Moving Average) и её модификации. Данная модель является комбинацией двух моделей: авторегрессии и скользящего среднего.

Основные проблемы при прогнозировании статистическими моделями:

- требование к рядам данных: для построения адекватной модели ARIMA требуется не менее 40 наблюдений, а для рядов, обладающих сезонностью – порядка 6–10 сезонов, что на практике не всегда возможно;
- неадаптивность моделей авторегрессии: при получении новых данных модель нужно периодически переоценивать;
- большие затраты ресурсов и времени на подбор параметров модели и предварительную обработку ряда [1].

Прогнозирование временных рядов моделями машинного обучения основывается на наличии набора данных для обучения и валидации (что уже является недостатком по отношению к статистическим моделям). Однако при этом данные модели являются более адаптивными (нет необходимости переобучаться при прогнозировании

нового ряда), а также многие из них не требуют сложной предварительной обработки входных данных. Градиентный бустинг обладает всеми этими преимуществами.

II. ПРИМЕНЕНИЕ ГРАДИЕНТНОГО БУСТИНГА

Градиентный бустинг – это техника машинного обучения для задач классификации и регрессии, которая строит модель предсказания в форме ансамбля слабых предсказывающих моделей. При этом обучение ансамбля проводится последовательно.

Бустинг, использующий деревья решений в качестве базовых алгоритмов, называется градиентным бустингом над решающими деревьями (Gradient Boosting on Decision Trees, GBDT). Он способен эффективно находить нелинейные зависимости в данных различной природы. Этим свойством обладают все алгоритмы, использующие деревья решений, однако именно GBDT обычно выигрывает в подавляющем большинстве задач. Благодаря этому он широко применяется во многих прикладных задачах (поисковом ранжировании, рекомендательных системах, таргетировании рекламы, прогнозировании погоды и многих других).

Рассмотрим решение задачи регрессии с помощью градиентного бустинга. В качестве оптимизируемой функции возьмём квадратичную функцию потерь:

$$L(x, y) = \frac{1}{2} \sum_{i=1}^N (y_i - a(x_i))^2.$$

Для решения будем строить композицию из K базовых алгоритмов:

$$a(x) = a_K(x) = b_1(x) + b_2(x) + \dots + b_K(x).$$

В качестве базовых алгоритмов выберем, как и условились, семейство решающих (регрессионных) деревьев некоторой фиксированной глубины.

Используя известные методы построения решающих деревьев, обучим алгоритм $b_1(x)$, который наилучшим образом приближает целевую

переменную:

$$b_1(x) = \operatorname{argmin}, L(y, b(x)).$$

Построенный алгоритм $b_1(x)$, скорее всего, работает не идеально. Вычислим, насколько сильно отличаются прогнозы этого дерева от истинных значений:

$$s_i^1 = y_i - b_1(x_i).$$

Теперь мы хотим скорректировать $b_1(x)$ с помощью $b_2(x)$. Для этого второе решающее дерево будет обучаться прогнозировать разности s_i^1 :

$$b_2(x) = \operatorname{argmin}, L(s^1, b(x)).$$

Ожидается, что композиция из двух таких моделей $a_2(x) = b_1(x) + b_2(x)$ станет более качественно предсказывать целевую переменную y .

Далее все действия повторяются до построения всей композиции. На k -ом шаге вычисляется разность между истинным ответом и текущим прогнозом композиции, затем k -й алгоритм обучается предсказывать эту разность, а затем обновляется вся композиция [2]:

$$s_i^k = y_i - a_{k-1}(x_i),$$

$$b_k(x) = \operatorname{argmin}, L(s^{k-1}, b(x)),$$

$$a_k(x) = a_{k-1}(x) + b_k(x).$$

Для обучения описанной модели и её применения в прогнозировании временных рядов необходимо иметь набор данных (т.е. некоторое множество рядов). При этом данных должно быть как можно больше, они должны содержать исчерпывающее количество видов зависимостей, а также в них должен отсутствовать выраженный дисбаланс.

Временной ряд можно представить как вектор признаков, где каждая его компонента - значение исследуемой величины в определенный промежуток времени. Однако для увеличения скорости обучения и работы алгоритма можно взять лишь несколько самых важных признаков (учитывающих недельную, месячную и годовую сезонность, а также текущий тренд). Таким образом, уменьшится размерность входных данных и, следовательно, их объём.

В качестве целевой переменной выступает следующее (неизвестное) значение временного ряда. Для прогнозирования первого неизвестного

значения в качестве входного параметра выступает вектор, полученный из первоначального ряда:

$$X^0 = (x_1, x_2, \dots, x_n).$$

Затем из данного вектора выбираются самые важные признаки (по порядковым номерам, определенным заранее):

$$x^0 = (x_i, x_j, \dots, x_k).$$

После получения первого неизвестного значения $y^0 = a(x^0)$, оно добавляется в конец исходного временного ряда:

$$X^1 = (x_1, x_2, \dots, x_n, y^0)$$

Затем уже из нового вектора X^1 выбираются самые важные признаки:

$$x^1 = (x_{i+1}, x_{j+1}, \dots, x_{k+1}).$$

Далее все действия повторяются до прогнозирования последнего p -го значения:

$$X^p = (x_1, x_2, \dots, x_n, y^0, y^1, \dots, y^{p-1})$$

$$x^p = (x_{i+p}, x_{j+p}, \dots, x_{k+p}).$$

$$y^p = a(x^p)$$

Таким образом, полученная модель на основе градиентного бустинга может принимать на вход любой временной ряд и прогнозировать следующие значения, основываясь на данных о других зависимостях из обучающей выборки (в отличие от статистических моделей, которые для прогноза используют лишь данные об исследуемом ряде).

III. ЗАКЛЮЧЕНИЕ

Описанный метод прогнозирования временных рядов обладает своими преимуществами и недостатками по отношению к рассмотренным статистическим методам. Выбор должен осуществляться исходя из имеющегося времени и ресурсов, а также из требований, предъявляемых к моделям и прогнозам.

СПИСОК ЛИТЕРАТУРЫ

1. Бизнес-прогнозирование / Д. Э. Ханк [и др.]. – пер. с англ. 7-е изд. М.: Вильямс, 2003. 506 с.
2. Градиентный бустинг [Электронный ресурс] / Школа Анализа Данных. – Режим доступа: <https://academy.yandex.ru/handbook/ml/article-gradientnyj-busting>. – Дата доступа: 17.10.2023.

АВТОМАТИЗАЦИЯ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ СТРУКТУРЫ МАТЕРИАЛОВ

Шиманский Н. А., Баглов А. В., Хорошко Л. С.
IDA Technologies

Кафедра физики твердого тела и нанотехнологий, физический факультет,
Белорусский государственный университет
Минск, Республика Беларусь

E-mail: nikita.shymanski@gmail.com, baglov@bsu.by, khoroshko@bsu.by

В данной работе рассмотрена возможность использования современных методов автоматизации и работы с большими данными, а также использования современных облачных технологий для решения прикладных задач материаловедения на примере одного из самых распространенных методов анализа структуры кристаллических материалов – дифракции рентгеновских лучей. Предложено программное решение, реализующее возможность эффективного масштабирования задач и глобализации базы данных. Приводятся примеры оптимизации решения задачи с применением разработанного решения.

ВВЕДЕНИЕ

Современная интеграция и глобализация международных исследовательских проектов делает актуальным создание общих баз данных, содержащих большое количество обработанных и верифицированных экспериментальных данных в виде, удобном для дальнейшего использования и обращения с ними. В настоящее время подобные базы данных, касающиеся структурных свойств материалов, могут быть как открытыми, так и предоставляемыми ограниченному кругу лиц в соответствии с приобретенной лицензией [1, 2], однако, конечная обработка результата поиска и сопоставление с собственным результатом осуществляется исследователем непосредственно, поскольку даже автоматизированные пакеты предполагают, во-первых, ручное сужение круга поиска, во-вторых, тщательное визуальное сопоставление и выбор окончательного решения.

С применением средств автоматизации, нейронных сетей и методов машинного обучения становится возможным не только предсказание строения и свойств наноматериалов, но и решение широкого ряда материаловедческих задач, а также снижение временных и трудовых затрат на эти процессы, что рассмотрено в данной работе на примере кейса рентгеновского дифракционного анализа материалов

I. АКТУАЛЬНОСТЬ И ЗАДАЧИ ИССЛЕДОВАНИЯ

Рассмотрим частный случай решения распространенной задачи исследовательского материаловедения на примере обработки результатов одного из базовых методов исследований свойств и структуры кристаллических материалов – дифракции рентгеновских лучей (*XRD*, от англ. «*X-Ray Diffraction*»), детальный анализ результатов которого может дать исчерпывающую информацию о строении материала, однако сопряжен с большим количеством специфических операций обработки [3-5]. Такие факторы как уровень шума на дифрактограммах, базовая линия шума,

наличие гало, разрывы данных и т.д. значительно увеличивают время на оценку и интерпретацию результатов. Для сложных дифрактограмм обработка и анализ результатов, особенно при отсутствии элементарной автоматизации, может превышать по времени проведение самого эксперимента *XRD*.

Также имеет место проблема оперативного предоставления доступа к результатам исследований заинтересованным исследователям, что на данный момент чаще всего реализуется созданием общих облачных хранилищ, пересылкой и-мейлов и т.п., но тем не менее, не отменяет многократной самостоятельной обработки результатов каждым исследователем в ручном режиме.

Современные объемы обрабатываемых данных, а также скорость их распространения все чаще требуют подходов *Big Data* (таких как *Data Warehouses / Data Lakes*), поскольку традиционные базы данных перестают демонстрировать устойчивую возможность обработки и анализа данных по мере роста их объема в хранилищах. Классические реляционные (а также *NoSQL*) базы демонстрируют высокие скорости записи и чтения данных, а также их чрезвычайную консистентность, но испытывают особые трудности работы с данными в тех случаях, когда хранилище данных возрастает до определенных пределов (терабайты) и должно быть распределено. В таких случаях операции параллельного чтения/записи могут значительно замедляться, что в конечном итоге приводит к перегрузке и полному отказу сервисов. Облачная платформа способна сохранять, каталогизировать, классифицировать и исследовать «на лету» практически неограниченные объемы дифрактограмм и иных спектроскопических данных (в том числе необработанных/неизвестных), полученных в ходе экспериментов.

II. ОПИСАНИЕ РЕАЛИЗАЦИИ

Платформа создаваемого решения представляет собой использование совокупности облачных сервисов глобального провайдера *Amazon Web Services (AWS)*, что подразумевает высокую доступность для пользователей во всех регионах планеты. Выбранная конфигурация в виде совокупности таких сервисов как *AWS Cognito*, *AWS DocumentDB*, *AWS S3 (Data Lake)*, *AWS Redshift / Spectrum*, *AWS SageMaker*, etc. гарантирует безотказный доступ к сервисам, позволяет горизонтально масштабировать нагрузки практически без ограничений, а также хранить и обрабатывать пользовательские данные, исчисляемые петабайтами.

Пользовательский интерфейс приложения позволяет проводить автоматизированную обработку дифрактограмм: удалить шум на графике, выровнять базовую линию шума, провести автоматический поиск пиков дифрактограммы и их позиций, а также провести аппроксимацию всех найденных или заданных пиков, и др.

III. ВЕРИФИКАЦИЯ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Методы математического моделирования, реализованные в приложении, позволяют обеспечить достаточно высокую точность в оценке распределения и позиционирования пиков *XRD*. Точность может быть дополнительно повышена увеличением количества итераций (циклов) при аппроксимации пика за счет дополнительного вычислительного ресурса.

По мере накопления материала в облачной базе данных достоверность прогнозов искусственного интеллекта растет, а в качестве стартового набора могут быть использованы как реальные дифрактограммы низкодефектных кристаллов, так и эталонные, сгенерированные специализированными научными программами (например, *Vesta* и др.). Полностью автоматическая оценка дифрактограмм, с точки зрения временных затрат на анализ полученных данных, значительно превосходит традиционную, а в рамках разрабатываемого проекта реализуется постоянное пополнение пользовательской базы данных и обучаемость моделей.

На примере постановки практической работы в учебной группе оценка затраченного времени для исследования типовой дифрактограммы показывает, что в среднем для полного анализа дифрактограммы в ручном режиме требуется от 2 часов, с использованием демо-версии лицензируемого платного пакета *Match!* от 20 мин, то платформе требуется в среднем 12–18 секунд для проведения автоматического анализа такой же дифрактограммы, при этом будет произведено устранение шумов и корректировка базовой линии.

IV. ВЫВОДЫ

Достоверность автоматизированного анализа с использованием нейросетей все еще является дискуссионным предметом в научном сообществе, но ее уровень, как и точность предсказаний, непрерывно растет по мере накопления банков данных *Data Warehouse / Data Lake* и совершенствования технологий в области *BigData* и *Machine Learning*, что открывает дальнейшие перспективы для дальнейшего углубления интеграции науки и информационных технологий. Методы обучения нейронных моделей при этом показывают перспективные результаты в области исследований и предсказания свойств наноматериалов [6]. Используемые вычислительные методологии не налагают строгих ограничений на источник и методики анализа материалов, которые могут быть автоматически обработаны, в частности, существует перспектива подобной автоматизации обработки результатов, полученных с помощью различных материаловедческих методов, включая Рамановское рассеяние (комбинационное рассеяние света), электронная Оже-спектроскопия, Резерфордское обратное рассеяние, люминесцентная спектроскопия и др.

1. Crystallography Open Database [Электронный ресурс]. – Режим доступа: <http://www.crystallography.net/cod/>. – Дата доступа: 16.10.2023.
2. Cambridge Structural Database System [Электронный ресурс]. – Режим доступа: <https://software.chem.ucla.edu/CSD/>. – Дата доступа: 16.10.2023.
3. Автоматизация обработки результатов исследования структуры и свойств наноматериалов / Н.А.Шиманский, А.В.Баглов, Л.С.Хорошко // BIG DATA и анализ высокого уровня = BIG DATA and Advanced Analytics : сб. науч. ст. IX Междунар. науч.-практ. конф. В 2 ч. Ч. 1 (Республика Беларусь, Минск, 17–18 мая 2023 года) / редкол. : В.А. Богдаш [и др.]. – Минск : БГУИР, 2023. – 296-300.
4. Облачные решения Big Data&Machine Learning в сфере автоматизации материаловедческих исследований / Л.С. Хорошко, А.В. Баглов // Система «Наука–Технологии–Инновации»: методология, опыт, перспективы: материалы Международной научно-практической конференции, Минск, 28-29 сентября 2023 г. / Под ред. В.В. Гончарова. –Мн.: Центр системного анализа и стратегических исследований НАН Беларуси, 2023. – в печати.
5. Трушин, В.Н. Рентгеновский фазовый анализ поликристаллических материалов / В.Н. Трушин, П.В. Андреев, М.А. Фаддеев // Электронное учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2012. – 89 с.
6. Использование методов глубокого обучения для решения научно-технических задач в области материаловедения / Л.С. Хорошко, А.В. Баглов // Система «Наука–Технологии–Инновации»: методология, опыт, перспективы: материалы Международной научно-практической конференции, Минск, 24-25 сентября 2020 г. / Под ред. В.В. Гончарова. –Мн.: Центр системного анализа и стратегических исследований НАН Беларуси, 2020. – С. 578–582.

ПРОИЗВОДСТВЕННАЯ ПРАКТИКА — КАК ЭТАП В ФОРМИРОВАНИИ СПЕЦИАЛИСТОВ В ОБЛАСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Ярмолик В. И., Боброва Т. С., Павловская Е. Р.

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

E-mail: {v.jarmolik, t.bobrova, pavlovskaya}@bsuir.by

Рассматриваются аспекты организации производственной (технологической) практики при подготовке специалистов в области информационных технологий.

ВВЕДЕНИЕ

Информационные технологии затрагивают сегодня практически все сферы общественной жизни и служат основой для экономического и социального развития государства. В соответствии с Декретом Президента Республики Беларусь «О развитии цифровой экономики» все большее внимание уделяется практико-ориентированному образованию [1].

Эффективная деятельность в профессии невозможна, если человек не видит себя в профессии, не представляет своей будущей деятельности, не имеет необходимой мотивации и не разделяет соответствующих ценностей. Но это все студент получает не автоматически при поступлении на выбранную специальность, а длительно, систематически и целенаправленно формируется в течение всего срока обучения в университете благодаря правильной организации процесса обучения и грамотным педагогам. Задача высшего учебного заведения состоит не в том, чтобы «наполнить» студента профессиональными знаниями, умениями и навыками, а в том, чтобы создавать условия, наиболее способствующие его личностному самоопределению и формированию профессиональной идентичности. Период профессионального обучения – это время развития самоидентификации, формирования личностной зрелости, освоения базовых ценностных представлений, определяющих выбранную специальность.

I. ОСНОВНЫЕ ПРОБЛЕМЫ ПРОВЕДЕНИЯ ПРАКТИКИ ПРИ ПОДГОТОВКЕ СОВРЕМЕННЫХ ИТ-СПЕЦИАЛИСТОВ

Значимую роль в формировании профессиональной идентичности играет производственная практика, на которой студент имеет возможность познакомиться со своей специальностью, скорректировать свой профессиональный образ, принять ценности и нормы профессии и реализовать их в реальной профессиональной ситуации. Производственная практика – незаменимая составляющая процесса обучения и формирования будущего специалиста. Именно на конкретном месте работы теоретический материал получает реальную оболочку, возникает соответствие между теорией и

практикой, происходит осознание того, чему и зачем учится студент. Производственная практика вносит в формирование профессионального самосознания личности больший вклад наравне с теоретической подготовкой. Производственная практика представляет собой двусторонний канал, позволяющий студентам принимать ценности и нормы профессии и реализовывать их в моделируемой ситуации.

Интеграция студентов в конкретную профессиональную деятельность во время производственной практики, а также ситуация реальной ответственности за свою деятельность позволяют им корректировать свой профессиональный образ, подкрепляя его требуемыми знаниями, умениями, навыками.

Практика способствует развитию самостоятельной работы студентов. В процессе прохождения практики студенты учатся самостоятельно отбирать и систематизировать информацию в рамках поставленных перед ними задач; применять полученные знания на практике; изучать технологию и оборудование, используемые в рамках конкретного производства; развивать навыки работы в коллективе; осуществлять самоконтроль.

Ответственность за организацию производственной практики возлагается на руководителей от ВУЗа и предприятия. Работа студента должна быть распланирована так, чтобы каждый день практики давал студенту возможность попробовать что-то новое. Время студента должно быть заполнено не скучным ожиданием окончания рабочего дня, а интересной работой, оттачивающей его знания, умения и навыки. Грамотное руководство производственной практикой влияет на сокращение периода адаптации выпускника к конкретным условиям работы [2].

Возросли требования работодателей к наличию опыта работы у выпускников, что отражает тенденцию, существующую на рынке труда в целом. Многие работодатели ждут от выпускника наличия опыта работы или совмещения учебы с работой на профильном предприятии, тогда как для трети работодателей наличие опыта работы не является обязательным. Названная причина является одной из основных трудностей, с кото-

рыми сталкивается выпускник на рынке труда по окончании обучения в вузе, когда предприятия и службы занятости в том числе, предпочитают иметь дело с соискателем вакансии, имеющим опыт трудовой деятельности. Решению этой проблемы должна способствовать гибкая организация производственной практики студентов, во время которой они могли бы иметь возможность получить опыт работы именно на том предприятии, в организации или учреждении, куда потенциально собираются трудоустроиться.

II. ПРЕДЛОЖЕНИЯ ПО УЛУЧШЕНИЮ ПРОВЕДЕНИЯ ПРАКТИКИ ПРИ ПОДГОТОВКЕ СОВРЕМЕННЫХ ИТ-СПЕЦИАЛИСТОВ

Практика реализует возможность овладения практическими умениями на основе теории, приобретенной в стенах университета. Для улучшения процесса прохождения практики работа преподавателя не должна ограничиваться только приемом готовых работ и выставлением оценок. Задача руководителя практики состоит в ответственном отношении к вопросам прохождения и выполнения программы практики в контакте с представителями предприятия. Совместные консультации студентов по процедурным вопросам, нормам поведения, грамотному составлению отчетных документов и сдаче готовых работ.

Компетентный подход руководителей способствует освоению программы практики при индивидуальной работе с каждым студентом. Поэтому важным является организация индивидуальной работы для пояснения сложных вопросов, понятным каждому студенту языком.

Для оформления документов и отчета по итогу практики необходимы более глубокие знания по делопроизводству и требованиям стандарта предприятия [3].

Среди технических проблем, связанных с практикой, стоит отметить недостаточность мест практики по специальностям, сокрытие практикантами фактов уклонения от прохождения практики, непрофессиональное поведение на месте практики, списывание отчетных работ.

Технические проблемы организации практики требуют проведения большей работы с предприятиями. Не все организации имеют специально выделенных специалистов для работы с практикантами, а ответственные специалисты на местах не всегда имеют достаточно времени и материальной заинтересованности для качественной работы со студентами. Поэтому требуется создание базы практики. База практики – это предприятие, учреждение или организация, в которую

для прохождения практики направляется группа студентов (студент) высшего учебного заведения.

В качестве баз для проведения производственной практики кафедрой высшего учебного заведения выбираются предприятия, соответствующие профилю подготовки специалистов в области информационных технологий, с которыми заключается договор о проведении производственной практики. Только во взаимном сотрудничестве университета и предприятия можно сформировать специалиста, отвечающего всем требованиям современного реального производства.

Таким образом, практика – это крайне важная часть образовательного процесса. Качество организации практики отражается на уровне приобретенных практических знаний и умений, и, в результате, на качестве работы выпускников ВУЗов. Для повышения качества организации практики необходимо учитывать проблемы организационно-методического характера, проблемы вовлеченности обучающихся и технические сложности. Их решение должно способствовать налаживанию системы рабочих и партнерских отношений между руководством образовательных организаций, руководителями практики, организациями-базами практики и обучающимися, чтобы готовить по-настоящему эффективных специалистов.

III. ЗАКЛЮЧЕНИЕ

Подводя итог, можно сказать, что производственная практика выполняет важнейшие функции в системе профессиональной подготовки студентов: Обучающую – актуализация, углубление и расширение теоретических знаний, их применение в решении конкретных ситуационных задач; формирование навыков, умений. Развивающую – развитие познавательной, творческой активности будущих специалистов, развитие мышления; коммуникативные и психологические способности. Воспитывающую – формирование социально активной личности будущего специалиста, устойчивого интереса, любви к профессии.

СПИСОК ЛИТЕРАТУРЫ

1. О развитии цифровой экономики : Декрет Президента Республики Беларусь от 21 декабря 2017 г. № 8 [Электронный ресурс]. – Режим доступа: <https://president.gov.by/ru/documents/dekret-8-ot-21-dekabrja-2017-g-17716>.
2. Программа по производственной (технологической) практике для специальности 1-53 01 02 «Автоматизированные системы обработки информации» УД-2-23/Пр. от 24.10.2017.
3. СТП 01-2017, Стандарт предприятия. Дипломные проекты (работы). Общие требования. – Минск : БГУИР, 2017. – 169 с.

НОВЫЙ АЛГОРИТМ ВЫБОРА ЛИДЕРА ПРИ АВТОНОМНОМ УПРАВЛЕНИИ РОЕМ БПЛА

Ярмолик В. И., Боброва Т. С., Ючков А. К.

Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: {v.jarmolik, t.bobrova }@bsuir.by

Рассматривается новый алгоритм поиска ведущего дрона (лидера) при децентрализованном подходе к управлению роем БПЛА.

ВВЕДЕНИЕ

В последние годы развитие беспилотных летательных аппаратов (БПЛА) достигло выдающихся результатов в различных областях военного, коммерческого и гражданского применения. Эффективность применения беспилотников значительно увеличивается при объединении их в группу, получившую название рой.

Каждый дрон управляется собственной автоматикой, а поведением роя может управлять программа с элементами искусственного интеллекта или один (несколько) операторов. Архитектура связи роя определяет, как осуществляется обмен информацией между дронами и между роем беспилотных летательных аппаратов и пунктом управления, находящимся на удаленном расстоянии.

I. ВАРИАНТЫ ИСПОЛЬЗОВАНИЯ РОЯ БПЛА

В течение продолжительного времени БПЛА использовались, в основном, для военных целей. Их применяли для поиска и уничтожения различных объектов и огневой поддержке с воздуха сухопутных и морских подразделений (войск).

В настоящее время на смену дистанционно пилотируемым аппаратам приходят автономные аппараты и рой БПЛА, которые могут самостоятельно выполнять поставленную задачу в автоматизированном режиме при минимальном вмешательстве человека-оператора. [1]

В зависимости от типа управления беспилотные летательные аппараты подразделяют на неуправляемые (такие, как метеорологические зонды), дистанционно пилотируемые и автономные.

Рой имеет ряд преимуществ, по сравнению с одиночными БПЛА, таких как живучесть, масштабируемость, скорость полета, автономность, стоимость, коммуникационные потребности, эффективная отражающая поверхность:

- более ускоренное получение результатов в различных областях применения;
- снижение потерь от аварий;
- уменьшение общей стоимости БПЛА;
- улучшение результатов обследуемого объекта за счет разных углов зрения БПЛА;

- улучшение точности позиционирования каждого дрона в рое за счет взаимного позиционирования.

II. АЛГОРИТМЫ УПРАВЛЕНИЯ РОЕМ БПЛА

Архитектура использования БПЛА включает либо ручное управление, с помощью дистанционного управления, либо автономное с помощью процессоров, установленных на беспилотных летательных аппаратах. Эти способы управления требуют наличие канала связи. Высокоэффективная и надежная связь обеспечивает возможность выполнения поставленных рою БПЛА задач, стратегию координации и совместимости, механизмы управления, безопасность, последовательность планирования миссии и многое другое. При централизованных стратегиях управления каждому БПЛА требуется канал связи с пропускной способностью в несколько Мбит/с для передачи изображений и другой информации об окружающей среде. Рассмотрим структуру обмена данными для роя БПЛА как процесс их взаимодействия для выполнения полученной цели (задачи). На начальном этапе становления такой структуры как рой БПЛА, применялась центральная станция, достаточно мощная для связи со всем роем БПЛА, так называемый централизованный подход, который позволил уйти от одиночного дрона к архитектуре взаимодействующих летательных аппаратов. Постепенно начал появляться децентрализованный подход, который дает возможность создать более сложную структуру и организацию роя, по снижающей его зависимость от центральных станций. [2]

Возможны различные варианты реализации децентрализованного подхода к управлению БПЛА. Рассмотрим несколько широко распространенных вариантов и укажем их преимущества и недостатки.

В случаях, когда группа БПЛА используется для задач, которые не подразумевают их потенциальное уничтожение вследствие обнаружения – задачи наблюдения за инфраструктурой, спасательных миссий и другие – чаще всего используется связь каждого отдельного БПЛА с множеством других, а также лидера группы с командным центром. В таких заданиях прак-

тически не возникает необходимости изменять лидера группы, следовательно, не нужно использовать никаких алгоритмов для поиска нового лидера при повреждении основного. Также стоит отметить, что не используется никакая внешняя инфраструктура – БПЛА связываются друг с другом при помощи встроенных модулей, которые позволяют обмениваться данными на ограниченном расстоянии, что также ограничивает гибкость управляемости роя.

Если группа БПЛА используется при наличии инфраструктуры, которая обладает достаточной пропускной способностью – например, сотовой сети 4 поколения и выше, и уничтожение дронов также не подразумевается, то возможна простая реализация связи каждого дрона с каждым и одновременно с командным центром. Данный подход не является полностью централизованным, так как практически всегда имеется связь не с одной центральной станцией, а с несколькими объектами инфраструктуры связи. В таком случае отпадает необходимость самого наличия лидера, а гибкость управления увеличивается – расстояние между отдельными БПЛА перестает иметь значение.

В случаях, когда нельзя полагаться на инфраструктуру, а также в случаях, где возможно уничтожение отдельных БПЛА, необходим алгоритм выбора лидера роя. Прежде чем рассмотреть возможные алгоритмы, опишем процесс опознания лидера в целом – он происходит, например, непосредственно при запуске роя БПЛА. Сначала первоначальный лидер передает сигнал всем дронам, которые находятся в его радиусе. Затем каждый из получивших сигнал дронов отправляет подтверждение назад лидеру и повторяют его сигнал. Процесс повторяется рекурсивно. В итоге получаем одного лидера, некоторое количество связанных дронов, и остальные, которые связаны либо с одним связным дроном, либо непосредственно с лидером.

Алгоритмы, связанные с переназначением лидера, могут использовать различные подходы к определению оптимального кандидата. Эффективность алгоритма можно оценить по скорости принятия решения, среднему периоду между переназначениями и задержкой соединения.

В качестве простейшего алгоритма выбора лидера можно предложить следующий: последовательно дадим номер каждому из дронов. Функционирующий дрон с минимальным номером прием за лидера. Преимуществом такого алгоритма является простота реализации, а также скорость принятия решения – так, например, если любой из дронов теряет соединение с лидером, то он сразу же принимает за лидера дрон со следующим номером. Однако если будут потеряны сразу несколько дронов с номерами, следующими за

лидером, то ожидание отклика каждого из них оставит рой без лидера на достаточно большое время.

Таким образом, необходимо использовать дополнительные параметры для выбора нового лидера. В качестве основных критериев, оценивающих оптимальность кандидата, стоит выделить количество соединений – чем больше, тем меньше будет задержка, а также заряд батареи – влияет на частоту смены лидера.

На основе данных параметров предложим следующий алгоритм: если соединение с текущим лидером потеряно, то каждый из БПЛА передает свой сигнал, содержащий количество собственных связей и заряд батареи. После этого, каждый из получивших сигнал БПЛА сравнивает данные параметры: если полученный сигнал имеет большее количество связей, либо равное при большем уровне заряда, то данный БПЛА признает источник сигнала лидером и дублирует его сигнал. Так происходит, пока весь рой не признает определенный дрон своим лидером.

Данный алгоритм имеет следующие преимущества: он является достаточно простым, позволяет передавать небольшое количество информации во время поиска, не оставляет рой без лидера на продолжительное время вне зависимости от порядка потери БПЛА. Его также возможно оптимизировать – к примеру, введением коэффициентов для количества связей и заряда батареи в зависимости от типа выполняемой задачи, а также, например, учитывать расположение БПЛА в пространстве – если определенная часть роя чаще теряет дронов, то это уменьшит вероятность выбора лидера из данной группы.

ЗАКЛЮЧЕНИЕ

В работе рассмотрены эффективность применения беспилотников, варианты использования роя, а также варианты выбора лидера при автономном управлении роем БПЛА. Описана связь внутри роя, а также, как осуществляется обмен информацией между дронами и между роем беспилотных летательных аппаратов и пунктом управления, находящимся на удаленном расстоянии.

1. Довгаль, В.А. Анализ систем коммуникационного взаимодействия дронов, выполняющих поисковую миссию в составе группы / В.А. Довгаль, Д.В. Довгаль // Майкоп, «Вестник АГУ». 4 (271), 2020. – стр.87-94..
2. Chen X, Tang J, Lao S. Review of Unmanned Aerial Vehicle Swarm Communication Architectures and Routing Protocols // Applied Sciences. Vol. 10, Iss. 10. 2020. – 3661 p.
3. Мунасыпов Р.А., Муслимов Т.З. Групповое управление беспилотными летательными аппаратами на основе метода пространства относительных состояний // Мехатроника, автоматизация, управление. 2018. № 2. С. 120–125.

IMAGE ENHANCEMENT ALGORITHMS

Yuan H., Gourinovitch A. B.
Faculty of Information Technology and Management,
Belarusian State University of Informatics and Radio Electronics
Minsk, Belarus
E-mail: 1846002512@qq.com, gurinovich@bsuir.by

Image enhancement is a process of improving image quality and enhancing image features by applying different methods and techniques. This article will introduce some common image enhancement methods and their principle formulas, and describe the final effects they produce.

INTRODUCTION

Image enhancement has wide applications in fields such as computer vision, image processing, and computer graphics. Image enhancement improves image visualization and analysis by improving aspects such as image contrast, clarity, color saturation, and detail. Image enhancement is widely used in many fields, including computer vision, medical imaging, remote sensing images, security surveillance, and digital art. It can improve the visual quality of images, enhance image features and information, and make images more suitable for subsequent analysis, processing and applications. The following are some common image enhancement methods and how they work.

I. HISTOGRAM EQUALIZATION

Histogram equalization redistributes the pixel values of the image so that the histogram of the image is evenly distributed throughout the entire grayscale range, thereby enhancing the contrast of the image. The following is its principle formula:

$$J(x, y) = (L - 1) * CDF(I(x, y))$$

Let the original image be I , the equalized image be J , and the gray level of the image be L (usually 256). Among them, CDF represents the cumulative distribution function, which calculates the cumulative probability of the occurrence of pixels at each gray level.

Histogram equalization can enhance the contrast of the image and make the details in the image more obvious. Dark areas will become lighter, and bright areas will become darker, resulting in a more even distribution of brightness throughout the image.

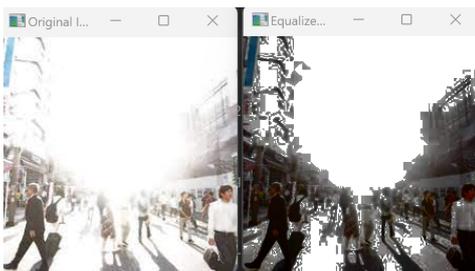


Figure 1 – Histogram equalization

The above figure is a comparison picture of pictures after histogram equalization. It should be noted that histogram equalization is a global operation and does not consider the local structure of the image. In some images, histogram equalization may introduce noise or enhance it. Therefore, before applying histogram equalization, you may want to consider using other image enhancement techniques or performing local contrast enhancement to process image details in specific areas.

In general, histogram equalization is a commonly used image enhancement technique that enhances the contrast and visual effects of an image by redistributing its pixel values. It is widely used in fields such as computer vision, image processing, and computer graphics.

II. GAUSSIAN FILTER

Gaussian Filtering is a commonly used image processing technique used to smooth images, reduce noise, and extract image features. It is based on the concept of Gaussian function and is implemented by performing a convolution operation on the image.

The Gaussian function is a commonly used mathematical function, also known as the normal distribution function. It has the shape of a bell curve, with a peak value and a standard deviation parameter. Gaussian functions are widely used in image processing to blur images and reduce noise.

Gaussian filtering is very common in real-time image processing, such as real-time object detection and tracking in the field of computer vision. Due to the linear and local nature of Gaussian filtering, the filtering operation can be accelerated by using fast convolution algorithms such as Gaussian pyramid.

The following is its principle formula:

$$J(x, y) = G * I(x, y)$$

Among them, let the original image be I , the Gaussian filter be G , and the filtered image be J . $*$ represents the convolution operation, and G represents the Gaussian filter template.

Gaussian filtering can smooth the image and remove noise and small details, making the image clearer. It reduces high-frequency components in the image by blurring the pixels in the image.

The following is a comparison picture of pictures after Gaussian filter:

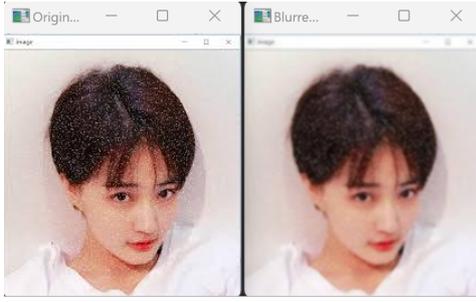


Figure 2 – Gaussian filter

It should be noted that the selection of Gaussian filter should be adjusted according to the specific application requirements. Smaller filter sizes preserve more detail, but may not reduce noise enough. Larger filter sizes smooth the image more strongly, but may result in loss of detail. Therefore, for some images, it may be necessary to select different filters or parameters according to specific needs, or to combine other image processing techniques to achieve better results.

III. SHARPENING (LAPLACIAN SHARPENING)

Laplacian sharpening is widely used in image processing and computer vision, and is often used to enhance the edges, details, and texture of images. It can make images clearer and more distinct, improving the visual quality and recognition performance of images.

By enhancing the high-frequency components in the image, the edges and details of the image are more prominent. The following is its principle formula:

$$J(x, y) = I(x, y) + k * L(x, y)$$

Among them, let the original image be I , the Laplacian filter be L , and the filtered image be J . k is the gain factor, and $L(x, y)$ represents the result of filtering the image by the Laplacian filter.

Laplacian sharpening enhances edges and details in an image, making them stand out more. It does this by adding high-frequency components to the image, making edges sharper and details clearer.

The following is a comparison picture of pictures after Laplacian filter:



Figure 3 – Laplacian filter

No matter which type of image you choose, you can enhance the sharpening effect by adjusting

the parameters of the Laplacian filter or combining it with other image enhancement techniques. Please note that when performing image processing, it should be taken to maintain the naturalness and authenticity of the image and avoid excessive enhancement or distortion.

IV. WAVELET TRANSFORM

Wavelet transform analyzes the frequency and spatial information of the image by converting the image into the wavelet domain, and enhances or suppresses the wavelet coefficients of different frequencies according to requirements.

Wavelet transform can capture the multi-scale details and structure of the image. By enhancing or suppressing the wavelet coefficients of different frequencies, the image details can be enhanced or removed.

The following is a comparison picture of pictures after Wavelet transform:

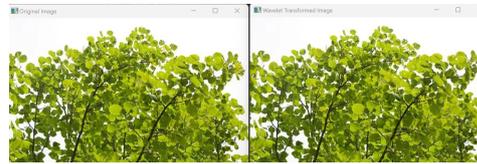


Figure 4 – Wavelet transform

Note that the effect of the wavelet transform also depends on the chosen wavelet function and transform parameters. You can try using different wavelet functions (such as haar, db2, db4, etc.) and adjust the transformation parameters to get the best results. Different wavelet functions and scale parameters are suitable for different types of signals and images. In addition, when displaying the wavelet transformed image, the contrast and brightness can be appropriately adjusted to enhance the visualization of the effect.

V. CONCLUSION

Image enhancement is the process of improving image quality and performance by applying various methods and techniques. This study introduces some common image enhancement techniques and models. The end effects that they produce are described. The methods are analyzed and it is shown under what conditions they are most effective.

REFERENCES

1. Xiaolin Chen, Research on Image Quality Evaluation Technology Based on Visual Features/ Shang Hai, 2012, Page 12-13.
2. Yanmei Li, Related Technology and Application Research of Image Enhancement /Chen Du, 2013,Page 15-20.

BLOCK METHODS OF MULTIDIMENSIONAL-MATRIX DATA ANALYSIS

Mukha V. S.

Department of Information Technologies of Automated Systems,
Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
E-mail: mukha@bsuir.by

The paper is devoted to the substantiation and the software implementation of the block method for the solving of the multidimensional-matrix problems. For this purpose, the notion of the cell is introduced as the ordered set of the multidimensional matrices. The two-dimensional block matrices are considered along with the cells as the multidimensional-matrix block constructions. It is shown that the multidimensional-matrix problem formulated in terms of the cells is transformed in the natural way to the block two-dimensional problem. As an application, the problem of the estimation of the coefficients of the multidimensional-matrix polynomial regression function is considered. This problem is reduced to the numerical solution of the system of the multidimensional-matrix linear algebraic equations. The two methods to the solution of the system are developed: the block Gauss method and the block method of the matrix inverse. Both of them are realized programmatically as standard functions.

INTRODUCTION

A block matrix is the ordinary (two-dimensional) matrix consisting of other ordinary matrices which are called the elements of the block matrix [1-5]. It is possible to perform the operations on block matrices similar to operations on the ordinary matrices. This circumstance led to the emergence of the block methods for solving the matrix problems. It is supposed that the matrices of the initial problem divide into the blocks, after that, the block solution methods are used. At the same time, such problems as choosing the number and nature of the blocks and the developing of the divide algorithm occur. In addition, the benefits of the block methods compared with the usual methods are not determined explicitly. In connection with these disadvantages, the block methods are not widely used in practice. As a consequence, there are no standard programs for the block matrix methods in the programming systems.

The situation changes radically, if the problem is generated initially in block form, i.e. if it is known that the matrices consist of the given blocks. The block approach is becoming the unique right approach in this case. It turned out, that the block form problems are generated in framework of the multidimensional-matrix approach. The purpose of this report is the statement of the natural connection between the multidimensional-matrix mathematical approach and block methods on the instance of the multidimensional-matrix regression analysis.

I. CELLS AND BLOCK MATRICES

Let A_1, A_2, \dots, A_m be q_1, q_2, \dots, q_m - dimensionality matrices respectively [6,7]. Let us call this ordered by one index $i = 1, 2, \dots, m$ set as one-dimensional cell, in accordance with the data type "cell array" of the Matlab programming system [8]. It is possible to order the set of the multidimensional matrices by the set of indexes. The set of

the multidimensional matrices ordered by p indices i_1, i_2, \dots, i_p is denoted $A = \{A_{i_1, i_2, \dots, i_p}\}$, $i_\alpha = 1, 2, \dots, m_\alpha$, $\alpha = 1, 2, \dots, p$, and is called p - dimensional cell.

The multiplication of the cells is determined by analogy with the multiplication of the multidimensional matrices [6, 7]. If $A = \{A_{i_1, i_2, \dots, i_p}\} = \{A_{l, s, c}\}$ is the p -dimensional cell, where $l = (l_1, l_2, \dots, l_\kappa)$, $s = (s_1, s_2, \dots, s_\lambda)$, $c = (c_1, c_2, \dots, c_\mu)$ are multi-indices, $\kappa + \lambda + \mu = p$, and $B = \{A_{j_1, j_2, \dots, j_q}\} = \{B_{c, s, m}\}$ is the q -dimensional cell, where $m = (m_1, m_2, \dots, m_\nu)$, $\lambda + \mu + \nu = q$, then the $(\kappa + \lambda + \nu)$ -dimensional cell $D = \{D_{s, c, m}\}$ is called the (λ, μ) -rolled product of the cells A and B , if its elements are defined by the expression $D_{l, s, m} = \sum_c^{\lambda, \mu} (A_{l, s, c} B_{c, s, m})$. We will denote the (λ, μ) -rolled product of the cells A and B as $\lambda, \mu \{AB\}$:

$$D = \lambda, \mu \{AB\} = \left\{ \sum_c^{\lambda, \mu} (A_{l, s, c} B_{c, s, m}) \right\} = \{D_{l, s, m}\}$$

We will consider the two-dimensional block matrices along with the cells. We will call them simply block matrices. It is advisable to transform the solution of the multidimensional-matrix problems to the solution of the two-dimensional-matrix problems due to the fact that the last problems are well developed. Such a transformation is performed on the base of the theorem on the associated matrices [6, 7] by replacing the multidimensional matrices with the associated with them two-dimensional ones. The received two-dimensional-matrix solution is then transformed back into the multidimensional-matrix solution. The forming of the associated matrices and the back transformation are performed on the program level. So the multidimensional-matrix problem formulated in terms of the cells creates the two-dimensional block problem.

II. EQUATION FOR THE COEFFICIENTS OF THE

It is assumed that the hypothetical regression function $y = \phi(x)$ is polynomial of the degree m [9, 11]:

$$y = \phi(x) = \sum_{k=0}^m {}^{0,kq}C_{p,kq}x^k = \sum_{k=0}^m {}^{0,kq}(x^k C_{kq,p}),$$

where x is the q -dimensional matrix input variable, y is the p -dimensional-matrix output variable, $x^k = {}^{0,0}x^k$ is the $(0,0)$ -rolled k -th degree of the matrix x , $C_{(p,kq)}$ are the $(p+kq)$ -dimensional matrices of the coefficients,

$$C_{p,kq} = (c_{i,\bar{j}_k}), i = (i_1, \dots, i_p), \bar{j}_k = (j_1, \dots, j_k),$$

satisfied the conditionals

$$C_{(p,kq)} = C_{(kq,p)}^{H_{p+kq,kq}}, C_{(kq,p)} = C_{(p,kq)}^{B_{p+kq,kq}},$$

$H_{p+kq,kq}$ and $B_{p+kq,kq}$ are transpose substitutions of the type "back" and "forward" respectively [6], and ${}^{0,kq}(C_{(p,kq)}x^k) = {}^{0,kq}(x^k C_{(kq,p)})$ are $(0,kq)$ -rolled productions of the multidimensional matrices. It is necessary to find the estimations $\hat{C}_{(p,0)}$, $\hat{C}_{(p,q)}$, ..., $\hat{C}_{(p,mq)}$ of the unknown coefficients $C_{(p,0)}$, $C_{(p,q)}$, ..., $C_{(p,mq)}$ on the base of the observations $(x_1, y_{o,1}), (x_2, y_{o,2}), \dots, (x_n, y_{o,n})$.

The following system of the linear algebraic equations for the coefficients $C_{(0,p)}$, $C_{(q,p)}$, ..., $C_{(mq,p)}$ is obtained in the work [10]:

$$\sum_{k=0}^m {}^{0,kq}(s_{x^{l+k}} C_{(kq,p)}) = s_{x^l y_o}, l = 0, 1, \dots, m \quad (1)$$

where

$$s_{x^{l+k}} = \frac{1}{n} \sum_{\mu=1}^n x_{\mu}^l x_{\mu}^k = \frac{1}{n} \sum_{\mu=1}^n x_{\mu}^{l+k},$$

$$s_{x^l y_o} = \frac{1}{n} \sum_{\mu=1}^n x_{\mu}^l y_{o,\mu}, y_{o,\mu} = (y_{o,i,\mu}), i = (i_1, \dots, i_p).$$

Let us note that the analytical (formula) solution to the system (1) is received for the constant, affine and square regression functions only [10].

The coefficients $s_{x^{l+k}}$ of the system (1) form the cell

$$A = \{a_{l,k}\}, a_{l,k} = s_{x^{l+k}}, l, k = 0, 1, \dots, m, \quad (2)$$

which elements $a_{l,k} = s_{x^{l+k}}$ are $(lq+kq)$ -dimensional matrices, the terms $s_{x^l y_o}$ form the one-dimensionality cell

$$B = \{b_l\}, b_l = s_{x^l y_o}, l = 0, 1, \dots, m, \quad (3)$$

which elements $b_l = s_{x^l y_o}$ are $(lq+p)$ -dimensional matrices, and the unknown variables $C_{(kq,p)}$ form

the one-dimensionality cell

$$Z = \{z_k\}, z_k = C_{(kq,p)}, k = 0, 1, \dots, m, \quad (4)$$

which elements $z_k = C_{(kq,p)}$ are $(kq+p)$ -dimensional matrices.

The multidimensional-matrix system (1) can be written in form of one equation in terms of cells A, B, Z (2), (3), (4):

$${}^{0,1}\{AZ\} = B, \quad (5)$$

i.e. as $(0,1)$ -rolled product of the cells, particularly, ${}^{0,1}\{AZ\} = \sum_{k=0}^m {}^{0,kq}(a_{l,k} z_k)$. We will solve the equation (5) and propose two algorithms for the solution.

III. BLOCK ALGORITHMS

The first algorithm to the solution of the equation (5) consists of use the block (generalized, in terminology of the work [1]) Gauss elimination algorithm. The equations (5) is equivalent, in accordance with the theorem on the associated matrices, to the following system of the equations in terms of associated two-dimensional matrices:

$$\sum_{k=0}^m ((\tilde{s}_{x^{l+k}})_{(lq,0,kq)} (\tilde{C}_{(kq,p)})_{(kq,0,p)}) = (\tilde{s}_{x^l y_o})_{(lq,0,p)}, \quad (6)$$

where $l = 0, 1, \dots, m$, $(\tilde{s}_{x^{l+k}})_{(lq,0,kq)}$ is the two-dimensional matrix $(lq,0,kq)$ -associated with the $(lq+kq)$ -dimensional matrix $s_{x^{l+k}}$, $(\tilde{C}_{(kq,p)})_{(kq,0,p)}$ is the two-dimensional matrix $(kq,0,p)$ -associated with the $(kq+p)$ -dimensional matrix $C_{(kq,p)}$, and $(\tilde{s}_{x^l y_o})_{(lq,0,p)}$ is the two-dimensional matrix $(lq,0,p)$ -associated with the $(lq+p)$ -dimensional matrix $s_{x^l y_o}$. The system of the equations (6) is represented by one equation

$${}^{0,1}[\tilde{A}\tilde{Z}] = \tilde{B} \quad (7)$$

with two-dimensional block matrix of the coefficients

$$\tilde{A} = [\tilde{a}_{l,k}], \tilde{a}_{l,k} = (\tilde{s}_{x^{l+k}})_{(lq,0,kq)}, l, k = 0, 1, \dots, m, \quad (8)$$

one-dimensional block matrix of the constant terms

$$\tilde{B} = [\tilde{b}_l], \tilde{b}_l = (\tilde{s}_{x^l y_o})_{(lq,0,p)}, l = 0, 1, \dots, m, \quad (9)$$

and one-dimensional block matrix of the unknown variables $\tilde{Z} = [\tilde{z}_k], \tilde{z}_k = (\tilde{C}_{(kq,p)})_{(kq,0,p)}, k = 0, 1, \dots, m$. The product ${}^{0,1}[\tilde{A}\tilde{Z}]$ in (7) means the block multiplication of the matrices. The received equation (7) is solved by Gauss elimination algorithm [1].

The second algorithm to the solution of the equation (7) uses the block inversion of the matrix \tilde{A} . The block solution of the equation (7) is $\tilde{Z} = {}^{0,1}[\tilde{A}^{-1}\tilde{B}]$, where \tilde{A}^{-1} is the matrix block-inverse to the matrix \tilde{A} .

The forming of the associate block matrices and back transfer to the cells are performed in accordance with the definition of the associate matrix [6, 7].

The block inversion algorithm is programmed as the block variant of the ordinary matrix inversion by the Gauss elimination [11]. The block multiplication follows the known rule [1]. All of the programs have the algorithmic generality, i.e. they allow solving the arbitrary order systems of the equations.

IV. CONCLUSION

It is showed that the block methods receive the new impulse for their application in the framework of the multidimensional-matrix mathematical approach. The problem of the numerical calculation of the coefficients of the multidimensional-matrix regression function is solved as the important application. The block methods can be used in other applications, for instance, in the problem of calculation of the coefficients of the orthonormal polynomials of the vector variable [7, 12].

REFERENCES

1. Gantmacher F.R. The theory of matrices. Moscow: Nauka, 1988, 552 p. In Russian.
2. Horn R.A., Jonson C.R. Matrix analysis. Cambridge university press, 2012, 662 p.
3. Korn G.A., Korn T.M. Mathematical Handbook for Scientists and Engineers, 2nd Edition. Dover, New York, 2000, 1152 p.
4. Ilyin V.A., Poznyak E.G. Linear algebra. Moscow, Nauka, 1974, 296 p. In Russian.
5. Tzon-Tzer Lu and Shiou Sheng-Hua. Inverses of 2x2 block matrices. Computers and Mathematics with Applications, 43, 2002, Pp. 119-129.
6. Sokolov N.P. Introduction to the theory of multidimensional matrices. Kiev, Naukova dumka, 1972, 176 p. In Russian.
7. Mukha V.S. Analysis of multidimensional data. Minsk, Technoprint, 2004, 368 p. In Russian.
8. Ketkov Yu.L., Ketkov A.Yu., Schults M.M. Matlab 6.x: programming of the numerical methods. SPb., BXV-Petersburg, 2004, 672 p. In Russian.
9. Mukha V.S. Analysis of multidimensional data. Laboratory manual. Minsk, BSUIR, 2018, 64 p. In Russian.
10. Mukha V.S. The best polynomial multidimensional-matrix regression. Cybernetics and Systems Analysis. V. 43, Issue 3, 2007, Pp. 427-432.
11. McCracken D.D., Dorn W.S. Numerical methods and Fortran programming. Moscow, Mir, 1978, 584 p. In Russian.
12. Mukha V.S. The system of the polynomials orthogonal with the discrete weight. Pro-ceedings of the National Academy of Sciences of Belarus. Physics and Mathematics Series. 2004, No 1, Pp. 69-73. In Russian.
13. Mukha V.S., Kako N.F. Model-based optimization of the regression objects // Information Technologies and Systems 2021 (ITS 2021). Proceeding of the International Conference (BSUIR, Minsk, Belarus, 24th November 2021). – Pp. 234–235.
14. Sokolov N.P. Introduction to the Theory of Multidimensional Matrices. Kiev, Naukova dumka, 1972. 176 p. In Russian.
15. Mukha V.S. Analysis of Multidimensional Data. Minsk, Technoprint, 2004. 368 p. In Russian.
16. Cichocki A., Lee N., Oseledets I., Phan A.-H., Zhao Q. and Mandic D.P. (2016). Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 1. Low-Rank Tensor Decompositions. Foundations and Trends in Machine Learning: Vol. 9: No. 4–5, pp. 249–429.
17. Ilyin V.A., Poznyak E.G. Linear algebra. Moscow.: Nauka, 1974. 296 p. In Russian.
18. Young, Eutiquio C. [1992] Vector and Tensor Analysis / Eutiquio C. Young – 2nd ed., rev. and expanded. – Marcel Dekker, Inc. – 511 p.
19. Mukha V.S. Multidimensional-matrix principal components method // Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series, 2013, no. 3, pp. 31–37. In Russian.

FEATURES AND CHARACTERISTICS OF ARTIFICIAL INTELLIGENCE SYSTEMS

Umaralizoda R. Sh., Kosimov A. A., Kamolidinov J. B.
Tajik Technical University named after academician M.S. Osimi
Dushanbe, Tajikistan
E-mail: rajab2000@yandex.ru

Artificial intelligence refers to the field of informatics, one of the tasks of which is to simulate human intelligence in an artificial way - with the help of a computer.

Artificial intelligence is a very promising field of research, the development of which is determined by achievements in the computer field. It can be assumed that the spread of computers and cybernetic devices, their use for human needs will determine the standard of living in the next century. The high productivity of new technologies largely depends on the use of artificial intelligence in them.

The most important practical result is the creation of expert systems – programs that simulate the work of a human expert in a certain subject area.

Under the control of the expert system, the computer, using information about the subject field, makes logical conclusions and composes answers to questions as a person – an expert in his field – would do.

Any expert system consists of three main parts – a database, a knowledge base, and logical inference programs. The database contains information about concepts and objects of the subject field. Knowledge base – information about their behavior and methods of interaction. Analysis of a specific situation, logical conclusions, and compilation of answers to questions are performed by logical deduction programs. The work of these programs is based on the principles of human intelligence.

An expert system is a computer system that embodies the experience of an expert based on his knowledge in a certain field. Based on the processing of this knowledge, the expert system (ES) can give intelligent advice, make decisions at the level of a professional expert, and also, at the user's request, explain the solution process in case of finding a particular solution [2]. The main differences between data processing systems and expert systems based on knowledge processing are as follows:

1. At the output of the expert system, the user receives not a machine or a videogram, which is presented in tabular form, but an intellectual advice in the form of text.
2. ES is based on the technology of symbolic information processing, which is mostly presented in the form of rules.
3. The expert system has an architecture that also differs from the architecture of traditional data processing systems. The differences consist in the presence in the expert system of

such blocks in: knowledge base; explanations; accumulation of knowledge [3].

The knowledge base is a collection of information about the subject area for which an expert system is being developed [4].

For the system to function, the knowledge base must be filled with knowledge. For this, highly qualified specialists in the field for which the system is being developed are invited, they play the role of experts whose task is to describe all known knowledge for the functioning of the ES. Knowledge of the first and second types must be available in the knowledge base. Knowledge of the first kind is generally known facts, phenomena, regularities, which are recognized in this subject area and published. Knowledge of the second kind is a set of empirical rules and intuitive conclusions that specialists use when making decisions under conditions of uncertainty in the presence of incomplete and contradictory information. Information about this knowledge, as a rule, is not published [1].

The ES knowledge base mainly contains knowledge of the first kind, but there should also be knowledge of the second kind. If this knowledge is missing, it means a bad choice of experts (they do not know how to formulate their knowledge or vice versa: they do not want to do it in order to preserve the status of unique specialists).

All knowledge presented in the knowledge base is divided into intentional and extensional. Intentional, or abstract, knowledge is notional (conceptual) knowledge about the objects of the subject area and the connections between them.

Extensional (specific) knowledge is the quantitative characteristics of the intentional part of knowledge, that is, the ES database [3].

A block of decisions, necessary for searching and constructing logical solutions, which are issued to the user by ES. The actions of this block are similar to the reasoning of a human expert who evaluates a problem and offers its hypothetical solution. This block performs the functions of managing the process of finding solutions, that is, it determines the way and sequence of using various rules and procedures. Each ES should contain a certain number of such rules and procedures. The number of rules contained in an average ES exceeds 500, and for a large ES it can exceed 1000 [3].

Mostly the junction block in consists of two parts: a logical output block and a control block. The block of logical inference performs actions similar to the intellectual activity of a specialist when he makes a decision. The functions of this block are the construction of a logical conclusion based on the existing knowledge stored in the knowledge base.

The control unit manages the process of finding a solution, that is, it determines the sequence of using various rules and procedures for manipulating knowledge.

The block of explanations serves to issue, at the user's request, a sequence of logical conclusions and considerations, which were used by the system in the process of finding a solution. The presence of such a block in the ES makes it possible to use it not only for making decisions, but also for the learning process as an educational system.

The problem of explanations and substantiation of the correctness of reasoning is an important and difficult task. After all, despite the fact that the system contains the knowledge of experts and provides advice, the user is personally responsible for the decision made. There are still no legal acts that would determine the responsibility of experts for the knowledge provided to the system, as well as for decisions made as a result of consulting users with the ES, which contains the knowledge of a given expert or group of experts.

In order for ES to be able to quickly and qualitatively explain the correctness of his answers, as well as the adequacy of the questions asked, he records the course of his reasoning and their sequence in his working memory.

The fact of being able to get an explanation gives the user the illusion that the system is checking the compliance of the rules recorded in the knowledge base. Meanwhile, the EU explains its actions only by issuing information about the progress of the reasoning process.

Justification is the text that the user will get after using this rule, intending to get an explanation.

The issue of automated acquisition of knowledge by an expert system (updating, correcting and expanding ES knowledge through the ES learning process) is of increasing interest. The problem of learning can be reduced to the creation of new concepts and rules on the basis of existing ones, as well as connecting them to the knowledge base in such a way that there is no contradiction of knowledge.

Creating new knowledge in a fully automated way is a very problematic approach proposed by overenthusiastic artificial intelligence advocates. Any knowledge (especially new) requires a very thorough check, which can only be performed by experts.

However, expert systems currently have some limitations:

- They usually work only in narrowly defined problem domains, and their level of under-

standing of the environment in which they work is somewhat superficial.

- These and systems still do not have "common sense", as means they certainly cannot think about the problem in multiple ways or at multiple levels. They don't know when they don't know anything relevant to the problem; it's a means that they can try to solve problems even when their knowledge and way of thinking is clearly inappropriate, and they won't be able to communicate that fact to the human user.
- They cannot learn by themselves.
- There are performance issues with response times on many systems.
- Expert systems can be expensive and risky to develop. Even when specialized hardware and software are not required, distilling human experience, encoding it, and storing it in knowledge bases for use in expert systems can be time-consuming and labor-intensive. The skill required for this is not available in many cases.

Despite these limitations, many computer corporations have developed applications of expert systems, both experimental and operational, that are deployed in use today. They concluded that these systems represent strategic competitive technology [1].

Research directions of artificial intelligence systems:

- development of theoretical foundations for the creation and application of artificial intelligence systems for various purposes;
- theoretical foundations and applied problems of creating intellectualized robotic systems;
- modeling of human intellectual activity and its application in artificial intelligence systems;
- of means and systems of intellectualization of computer interfaces;
- development of intelligent control systems for autonomous robots and robotic complexes;
- creation and application of highly intelligent multimedia and hypermedia technologies and tools for artificial intelligence systems;
- creation of educational programs and virtual environments with elements of artificial intelligence;
- analysis, synthesis and modeling of neural networks, development of methods of their design, optimization and training;
- development of neurocomputer application technologies, applied systems based on neural networks [2].

An expert system is a system that contains information about concepts and objects of a certain subject area, information about their behavior and methods of interaction, as well as logical conclusion programs; it is intended to simulate the work of a

human expert in this field. The main characteristics of the expert system are as follows:

1. The expert system, as a rule, is limited to a certain subject area.
2. ES must be able to make decisions based on incomplete or inaccurate data.
3. ES should be able to explain their actions when solving the problem.

REFERENCES

1. Bondarev, V.N. Systems of artificial intelligence. – Sevastopol: Publishing House of SevNTU, 2002. – 615 p.
2. Zgurovsky, M.Z. System analysis: problems, methodology, applications: "Scientific book" project. – K.: Nauk dumka, 2005. – 743 p.
3. Perevozchikova, O.L. Fundamentals of system analysis of objects and processes of computerization: study guide. – K.: Publishing House "KM Academy", 2003. – 431 p.
4. Soroka, K.O. Fundamentals of systems theory and system analysis: Study guide. – Kharkiv: A.N. Tymchenko Publisher, 2005. – 286 p.

WEB SYSTEM FOR AUTOMATION AND DIGITALIZATION OF THE PROCESSES OF ORGANIZATION AND SUPPORT OF GRADUATE WORK DESIGN AT THE GRADUATING DEPARTMENT

Savenko A. G., Khotak S. A.

Institute of Information Technologies of the Belarusian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: savenko@bsuir.by

The paper presents a developed web system for automation and digitalization of the graduate work design process. The proposed solution simplifies all stages of the process, and also allows you to monitor the progress of thesis design through an electronic journal, right up to the defense of thesis projects.

INTRODUCTION

From the point of view of organization and support, the process of designing graduate work (PDGW) is a rather complex and time-consuming task due to the presence of a large number of different stages and a large number of participants in this process, especially in departments that graduate a large number of students annually. Often, at various stages of organizing and maintaining PDGW, participants in this process make various types of errors, primarily due to the large amount of data. Research into the processes of organizing and managing coursework and diploma projects in the educational process has revealed problems such as the decentralization of process management tools and information about the progress of processes [1], the frequent detachment of project topics from existing needs and tasks solved in the real sector of the economy [2]. In addition, the general trends of informatization and digitalization of education, which have their own advantages [3], should be reflected at this stage of the educational process. The use of ICT in the process of organizing PDGW also has a positive effect on improving the quality of the educational process [1].

I. PROPOSED SOLUTION

Taking into account many years of accumulated experience in supporting the process, the graduate department of information systems and technologies of the BSUIR has developed and implemented a web system for automating the processes of organizing and supporting the PDGW, which allows minimizing errors caused by the human factor, simplifying the tasks performed for all participants in the process (including students), concentrate all the necessary information in one place and increase the visibility of the results obtained. In addition to students, the PDGW involves graduate work supervisors (both staff members of the department and external specialists), technical consultants, normative inspectors, feasibility study consultants, members of working commissions, reviewers, process managers, secretaries of the state examination commission (SEC) and head of the department. The PDGW itself,

from an organizational point of view, consists of the following stages: choosing a topic for a diploma project, choosing a supervisor, completing and defending pre-diploma internship, checking a feasibility study, passing regulatory control, passing a working commission (pre-defense), reviewing and defending a diploma project at the SEC, and also drawing up reporting documentation and draft orders and instructions at all stages. All these stages are automated and digitalized to a certain extent through the use of a web application and are controlled by process manager. The developed web system has three access levels: “student”, “teacher” (includes supervisor, consultants, normative inspectors, etc.) and “administrator” with full functionality (process manager and head of department). To use the application, users must register and consent to the processing of their personal data. To verify system users, only those users who are listed in the system database can register: the student must be a final year student, and the teachers must be listed as employees of the department. Such addition to the system database can be done manually (for special cases) or through synchronization with other databases of systems used at the university (dean’s office, human resources department, etc.) by uploading spreadsheets. All registration processes, and then filling out all fields for entering any information, are accompanied by prompts that can be generated and edited from the system administrator’s personal account. After registration, a user of the “student” access level indicates the place of pre-diploma internship, in accordance with the concluded agreement (necessary for generating reporting documentation). Next, the student is asked to choose the topic of the thesis project and the thesis supervisor. To increase the practice-oriented nature of diploma projects, the student is also given the opportunity to propose his own topic for the diploma project, for example, related to his work, hobbies and scientific research, and then send this topic for approval and adjustment to the selected supervisor. Moreover, the student can choose as a supervisor either a full-time employee of the department from the system database (the maximum number of graduates for each teacher is synchronized with the planned workload accord-

ing to the teacher's individual plan), or propose an external supervisor. When selecting an external manager, the corresponding application is displayed in the personal account of the process manager, who carries out verification and, if necessary, approval. When a student selects a supervisor, a letter is sent to his email with a link to his personal account, where an application from the student is awaited. The supervisor sees information about the student and the selected topic of the diploma project. The supervisor can accept or reject the application and, if accepted, subsequently adjust the topic. While the supervisor has not accepted the application, the student can withdraw it and send a request to another supervisor, as well as edit the proposed topic. After accepting a request from a student, an electronic journal of the PDGW is displayed in the teacher's and the student's personal account. Moreover, the right to change various elements of the electronic journal has the appropriate persons responsible for the various stages of the PDGW (supervisor, technical consultant, normative inspector, economic consultant, chairman of the working commission, etc.). The student can change the choice of a date convenient for him to defend his thesis project (SEC) after passing the working commission. Since the manager displays the electronic journals of all his graduate students, various filters are provided in his personal account.

Студент: Фамилия2 Имя2 Отчество2

Информация о дипломном проекте:

Тема: Программный модуль мониторинга учета рабочего времени в организации по оказанию бытовых услуг

Место проведения практики: ООО "Компания2" Город: Минск

Практика: с 27.10.2023 по 29.11.2023 Оценка за практику: 8

Опроценки:

до 23.11.2023	до 04.12.2023	до 18.12.2023
25	45	

ТЭО:

Дата: 27.11.2023	Статус: Не готов
Дата: 01.12.2023	Статус: Готов

Нормоконтроль:

Дата: 15.12.2023	Статус: Частично готов
Дата: 22.12.2023	Статус: Готов

Рабочая комиссия №1:

Дата: 04.01.2024	с 12:00	по 14:00	Адрес: 803-7
------------------	---------	----------	--------------

ГЭК: ВЫБРАТЬ ДАТУ

Figure 1 – Example of an electronic journal

All appointments of responsible persons and reviewers are made from the personal account of the process manager (system administrator). Moreover, this process is also automated, for example, you can appoint a technical consultant for all students whose supervisor is a certain employee, or assign a normative inspector for all students of a given group, etc.

Савенко Андрей Геннадьевич (Админ) ДП: Test Выйти

Назначения: НОРМОКОНТРОЛЬ ТЭО РАБОЧЕ КОМИССИИ РЕЦЕНЗЕНТЫ

Нормоконтроль

Группа: 981061 Ответственный: Савенко Андрей Геннадьевич

Группа: 981063 Ответственный: Карпечин Ольга Юрьевна

Группа: 981064 Ответственный: Колоткин Иван Александрович

Figure 2 – Example of an electronic journal

Also, in the personal account of the process manager, the parameters of the electronic journal are configured (for example, performance standards and their gradation when students pass checkpoints (percentages), the date of passing checkpoints, the date of the defense (SEC) and the maximum number of projects heard on this date. An extremely important feature is the function of automatic generation of reporting documentation and draft orders for the process manager. It allows you to generate text documents according to the required pattern of the following documents: a draft order on pre-diploma practice, a report on the beginning of practice, a report on completion of practice, a draft order on the approval of the topics of diploma projects and supervisors. Documents are generated with one click.

II. CONCLUSION

The developed web system significantly simplifies and makes convenient the process of organizing, supporting and completing a diploma project for all participants in this business process, allows students to automatically offer students topics for diploma projects related to real projects, select and propose managers, minimizes errors caused by the human factor, which is especially important for a department that graduates a large number of students. The web system was introduced into the organization of the process at the Department of Information Systems and Technologies of the BSUIR and is hosted on the university server in the public domain.

BIBLIOGRAPHY

1. Mal'kovskaya M.K. Problemy diplomnogo i kursovogo proektirovaniya i puti ikh resheniya. [Electronic resource]. - Access mode: <https://emc21.ru/wp-content/uploads/2020/09/Statya-Malkovskaya.pdf>. Access date 10.10.2023.
2. Nazarov A.A. Organizatsiya diplomnogo proektirovaniya v sovremennykh usloviyakh / A.A. Nazarov // Problemy gosudarstvennoi attestatsii vypusknikov: materialy nauchno-metodicheskoi konferentsii. - Vladimir : VGU, 2006. - S. 17-18.
3. Karpekin, I. A. Preimushchestva i effektivnost' vnedreniya distantsionnoi formy obrazovaniya v obrazovatel'nyi protsess uchrezhdenii obrazovaniya lyubogo tipa / Karpekin I. A., Savenko A. G. // Distantsionnoe obucheniye – obrazovatel'naya sreda XXI veka : materialy XI Mezhdunarodnoi nauchno-metodicheskoi konferentsii. – Minsk : BGUIR, 2019. – S. 136–137.

THE REDUNDANT FUNCTIONS REDUCING ALGORITHM

Tang Yi, Zhao Di, Gourinovitch A.B.

Department of Information Technologies in Automated Systems,
Belarussian State University of Informatics and Radioelectronics

Minsk, Republic of Belarus

E-mail: tangyijcb@163.com, 3189124246@qq.com, gurinovich@bsuir.by

Deep convolutional neural networks (CNNs) have achieved remarkable success in computer vision tasks. However, the redundancy of features in CNNs gradually increases with the deepening of the network and the increase of parameters, leading to an increase in the storage and computational overhead of the model. To address this problem, this paper proposes a channel reconstruction convolutional approach to optimize the CNN model by reducing the redundancy of features and enhancing the representational ability of features.

INTRODUCTION

Deep Convolutional Neural Networks (CNNs) have become an important tool in the field of computer vision, and have achieved remarkable results in tasks such as image classification, target detection, and semantic segmentation.

CNNs are able to efficiently extract features from an image and classify or localize them through fully connected layers by means of multilayer convolution and pooling operations. However, as the network continues to evolve, the depth of the network and the number of parameters continue to increase, leading to an increase in feature redundancy in CNNs[1].

Feature redundancy refers to the presence of similar or duplicate feature representations in different layers of the network. This redundancy not only increases the storage and computational overhead of the model, but also may lead to overfitting problems and reduce the generalization ability and robustness of the model. In addition, redundant features may affect the convergence speed and training stability of the network, making it more difficult to optimize the model.

I. PROBLEM STATEMENT

Convolutional Neural Networks (CNNs) have achieved great success in areas such as image and speech, but the convolution operation itself may have some redundancy[2]. Shown below are some of the current challenges and difficulties.

Model interpretability: removing the redundancy of convolution may lead to feature reduction and distortion, making the decision-making process of the model more difficult to interpret.

Feature representation capability: convolutional neural networks have been successful in many fields, partly due to their good feature representation capability. However, removing redundancy may weaken the representation of the feature maps, thus affecting the performance of the model.

Computational complexity: convolutional operations are one of the key components in deep learning models, and removing redundancy may introduce additional computational complexity.

II. DESCRIPTION OF EXISTING METHODS

In terms of reducing feature redundancy, researchers have proposed a variety of methods, including techniques such as sparse coding, feature selection, and feature pruning. However, each method has some different disadvantages and advantages.

Sparse coding methods can reduce redundancy by limiting the sparsity of features and learning more compact feature representations. However, sparse coding methods have high computational complexity and the training time may be too long, especially for large-scale data and complex models. In addition, sparse coding methods are sensitive to the choice of hyperparameters and require extensive tuning experiments.

Feature selection methods can help reduce computational overhead and improve model interpretability by selecting the most representative subset of features. However, the design of selection criteria often needs to consider the characteristics of specific tasks and data, which may lead to the loss of some useful information. Selection criteria that are too strict or inaccurate may exclude some potentially valuable features, thereby affecting model performance. Furthermore, the feature selection process is also affected by correlations and dependencies between features.

Feature pruning methods remove redundant feature channels, which can reduce computational and storage overhead while simplifying the model and improving its efficiency. However, excessive pruning may result in the loss of useful features, and calculating the importance of feature channels requires additional computational and storage overhead.

Introducing the attention mechanism to dynamically select and weight different feature maps according to the task requirements. Using the residual connection to alleviate the problem of gradient vanishing and improve the training effect and generalization ability of the model. Combine multi-scale and multi-modal information to improve the feature perception ability of the model in complex situations, and finally improve the performance of the model on various tasks and data.

In order to overcome these problems, a new method called "parallel channel reconstruction convolution" was proposed to improve the efficiency, generalization ability, and robustness of the model.

III. METHOD DESCRIPTION

The parallel reconstruction convolution module is divided into two branches, the upper and lower. The structure of the proposed module is shown below in Figure 1.

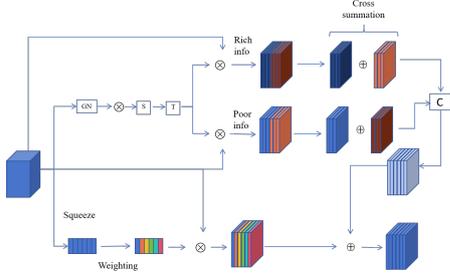


Figure 1 – The structure of parallel channel reconstruction convolution

The upper branch focuses on the spatial reconstruction of features. This branch consists of two parts: splitting and reconstruction. In the splitting part, first, a group normalization operation is performed and different groups are assigned weights to be multiplied with the group normalization result and finally a sigmoid activation function is used. The computation for assigning weights for group normalized features can be denoted as formula(1):

$$W_r = \frac{r_i}{\sum_{j=1}^C r_j}, i, j = 1, 2, \dots, C \quad (1)$$

Where W_r represents the normalized correlation weights mapped to range (0,1), C stands for number of channels.

The spatial pixel variance of each batch and channel is measured using the trainable parameters in the GN layer. Where richer spatial information reflects more variation in spatial pixels[4], contributing larger values r . Normalized correlation weights are obtained through equations indicating the importance of different feature maps. The results are classified according to a predefined threshold, and those greater than the threshold are considered to contain richer information, while others are considered to contain less information. Meanwhile, to reduce spatial redundancy, generate richer features, and save space, in the proposed reconstruction operation, instead of directly adding the information-rich features with the less informative ones, a cross-reconstruction operation is used to fully combine the weighted different informative features and enhance the information flow between them.

The lower branch focuses on the reconstruction of features over channels. Specifically, the SE layer is employed as a content-aware layer to re-weight

the mechanisms of each channel to create robust representations[5]. Make the network more sensitive to important features while suppressing irrelevant features. This goal is accomplished in two steps. First, the feature map is compressed using global average pooling to get a comprehensive view of each channel. The result of the compression operation is a feature vector of size n , where n denotes the number of channels. The second step is excitation, where this feature vector is fed through a two-layer feed-forward neural network, where the number of features is first reduced and then expanded to the original size n . This vector of size n now represents the weights of the original feature map, which can be used to scale each channel.

Finally, the results of the above two branches are connected to get the final output, which realizes the reconstruction of the input features.

IV. CONCLUSION

The channel reconstruction convolution method proposed in this paper is an effective way to optimize CNN models by reducing feature redundancy and enhancing feature representation capability to improve model performance. This method can significantly reduce feature redundancy while improving the performance and generalization ability of the model.

Future research can further explore the application of the channel reconstruction convolution method in the above tasks and incorporate more optimization strategies to further improve the performance of the model. For example, attention mechanisms can be combined to further optimize feature selection and weighting for better extraction of useful information[6]. Meanwhile, more complex network structures and model designs can be explored to improve the characterization of features and the performance of the model. Through continuous research and improvement, the channel reconstruction convolutional method is expected to play a greater role in the field of computer vision and bring more benefits to practical applications.

1. Qiu J, Chen C, Liu S, et al. Slimconv: Reducing channel redundancy in convolutional neural networks by features recombining[J]. IEEE Transactions on Image Processing, 2021, 30: 6434-6445.
2. Chen J A, Niu W, Ren B, et al. Survey: Exploiting data redundancy for optimization of deep learning[J]. ACM Computing Surveys, 2023, 55(10): 1-38.
3. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
4. Chen C F, Oh J, Fan Q, et al. SC-Conv: Sparse-complementary convolution for efficient model utilization on CNNs[C]//2018 IEEE International Symposium on Multimedia (ISM). IEEE, 2018: 97-100.
5. Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
6. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

DILATED CONVOLUTION AND SPATIAL PYRAMID FUSION IN THE IMAGE SEGMENTATION PROBLEM

Zhao Di, Tang Yi, Gourinovitch A.B.

Department of Information Technologies in Automated Systems,
Belarussian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus

E-mail: 3189124246@qq.com, tangyijcb@163.com, gurinovich@bsuir.by

Image segmentation is one of the important tasks in the computer vision, where the goal is to segment different regions in an image into semantically meaningful parts. However, due to the presence of target and contextual information at different scales in an image, traditional segmentation methods face the challenges of information loss and lack of accuracy when dealing with images at different scales. To address this problem, this study proposes an innovative approach that combines dilated convolution and spatial pyramid pooling to improve the processing power and accuracy of segmentation models for images of different scales.

INTRODUCTION

Image segmentation technology has important application value in medical image analysis, automatic driving, image editing and other fields. However, traditional image segmentation methods have certain limitations when dealing with medical images of different scales. This is because traditional methods usually use a fixed-size convolutional kernel to extract features, which cannot effectively capture the correlation between targets and contextual information at different scales, leading to performance degradation when processing images with multi-scale features, and facing the challenges of information loss and lack of accuracy. To address these problems, methods such as dilated convolution and spatial pyramid pooling are widely used in image segmentation.

To overcome these limitations, this article proposes an effective method that combines two key techniques: dilated convolution and spatial pyramid pooling. Dilated convolution captures multi-scale contextual information, while spatial pyramid pooling facilitates the fusion of multi-scale features. By integrating these techniques, the segmentation model can be enhanced to handle medical images at different scales, thus improving the accuracy and processing of multi-scale features.

I. PROBLEM STATEMENT

Medical images usually have complex anatomical structures and detailed information at different scales. In recent years, researchers have proposed some effective methods to improve the image segmentation method, but they still have limitation problems when dealing with images at different scales [1]. Several of these methods are described below:

Multiscale Pyramid: Constructing a multiscale pyramid allows obtaining feature representations at different scales for better handling of multiscale targets. However, this requires the introduction of multiple branches or convolution kernels at multiple scales in the network, increasing the number of parameters and computational complexity of the

network. In addition, when fusing features from different scales, the fusion strategy needs to be carefully designed to avoid confusion or conflict of information.

Residual connectivity: Residual connectivity can facilitate the transfer of information and the flow of gradients, which helps to improve the performance of image segmentation. However, when dealing with images at different scales, lower resolution features may interfere with higher resolution features, leading to inaccurate segmentation results.

Multi-scale fusion: Multi-scale fusion aims to combine feature information from different scales to improve the performance of the segmentation model. However, multiscale fusion may require additional computational and memory overheads, especially in feature-level fusion.

Dilated Convolution: By using a larger dilated rate, the sensory field can be expanded to capture a wider range of contextual information. However, a larger dilation rate may lead to a decrease in the resolution of the feature map, which may reduce the accuracy of details and boundaries.

II. METHOD DESCRIPTION

Firstly. The article uses dilated convolution to expand the sensory field of the convolution operation. In traditional convolutional operations, the size of each convolutional kernel is fixed and sliding computation is performed over a local region. However, this operation can only capture a limited range of contextual information, limiting the model's ability to understand a larger range of features. In contrast, dilated convolution increases the effective size of the convolution kernel by inserting fixed intervals of zeros within the convolution kernel. This interpolation operation is called dilated, and the dilation rate determines the size of the interpolated interval. A larger dilation rate effectively expands the receptive field, allowing the convolution operation to capture a wider range of contextual information, Figure 1 shows the schematic diagram of dilated convolution.

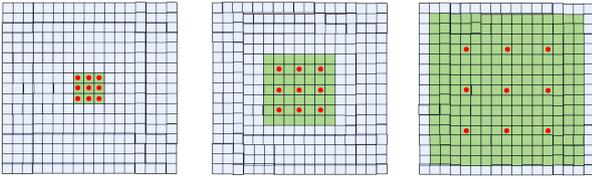


Figure 1 – The schematic diagram of dilated convolution

Figure 1 shows the following: Convolution results in a convolution kernel of the same size receiving a larger sensory field by introducing a dilation rate parameter. From left to right the different sensory fields are shown. When the dilation rate = 1, 2, and 4 respectively, then the size of the sensory field can be calculated from the size of the convolution kernel and the expansion rate. Assuming that the convolution kernel size is K and the expansion rate is R , the formula for the sensory field (F) is denoted by formula (1):

$$F = K + (K - 1)(R - 1) \quad (1)$$

By introducing dilated convolution, the sensory field can be increased without increasing network parameters. This is important for dealing with targets and scenes at different scales, since the relationships between features and structures at different scales require a larger range of contextual information to be accurately understood.

Secondly. For the CNN structure, the change of the image size mainly affects the fully connected layer (FC), so [2] adds the spatial pyramid pooling layer, between the convolutional layer (including convolution, Pooling, etc.) and the fully connected layer (FC). It allows the fully connected layer to the fixed number of features obtained and removes the restriction on the fixed size of the network. The spatial pyramid pooling module consists of multiple parallel pooling layers with different sense field sizes.

By transforming input images into fixed size outputs, spatial pyramid pooling can adapt to different sizes of input images. Its main idea is to divide the input image into grids of different sizes, carry out pooling operations on each grid, and then stitch all the pooling results together as the output. In this way, a fixed size output can be obtained regardless of the size of the input image. With pyramid pooling, it can aggregate features at different scales to capture information at different scales in the image.

Compared with the traditional fixed-length feature vectors, spatial pyramid pooling can better preserve the feature information at different scales and avoid information loss. The structure of spatial pyramid pooling is shown in Figure 2.

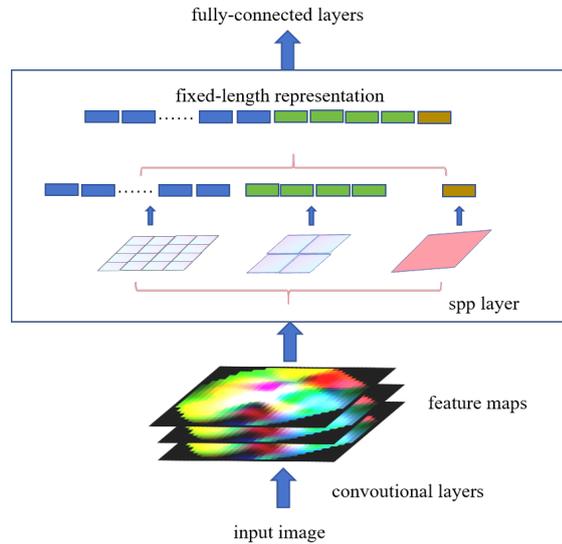


Figure 2 – The structure of spatial pyramid pooling

By performing pooling operations on feature maps at different scales, multiple feature representations at different scales can be obtained. In each pooling layer, the article performs average pooling or maximum pooling operations on the pooled feature maps to generate fixed-length feature vectors. These feature vectors represent feature information at different scales[3]. In this way, it avoid the problem of information loss that may result from using a single fixed-length feature vector to represent features at different scales. Finally, the article fuses these multi-scale feature vectors, which can be weighted and fused using a simple splicing operation or by introducing an attention mechanism. The fused feature vectors can be passed to subsequent classification or segmentation modules for final prediction.

III. CONCLUSION

The innovative approach is proposed to handle the task of image segmentation at different scales by combining dilated convolution and spatial pyramid pooling. It is able to efficiently capture contextual information at different scales and significantly improve the performance and robustness of the segmentation model. Further study can explore how to optimize the structure of dilated convolution and spatial pyramid pooling. The combination with other advanced techniques improves the model's efficiency. It is necessary to advance the development of image segmentation techniques and achieve better results in various practical applications.

1. Yu, F., & Koltun, V. (2016). Multi-Scale Context Aggregation by Dilated Convolutions. In International Conference on Learning Representations (ICLR).
2. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904-1916.
3. Chen, L., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. *ArXiv*, abs/1706.05587.

Содержание

Пленарные доклады	7
Голенков В. В., Гулякина Н. А., Шункевич Д. В. Ключевые проблемы и стратегические цели работ в области искусственного интеллекта	9
Якушкин С. И., Владимиров К. И., Пузикова В. В. RISC-V – процессорные технологии без ограничений	15
Секция «Системы управления»	17
Арбузов Г. Н., Сипович А. С., Шобик И. П. Выбор оптимального алгоритма обработки сигналов инерциального датчика из состава рабочего места балансировки роторов	19
Арбузов Г. Н., Сипович А. С., Шобик И. П. Сравнение алгоритмов обработки сигнала акселерометра используемого для определения неуравновешенности ротора в стенде балансировки	21
Ахапкина А. М., Способ С. П. Интеграция нейронных сетей в современные системы управления беспилотными автомобилями	23
Ахапкина А. М., Стежко И. К. Нейронные сети в автономном управлении: путь к безопасным и эффективным транспортным средствам	25
Гурская Я. С., Марков А. В. Проблемы автоматизации процессов производства молочной продукции	27
Ломонос Г. В., Захарьев В. А. Современные системы обнаружения вредоносных сообщений электронной почты	29
Маркова А. А., Кузнецов А. П. OLAP-технологии и визуализация информационных ресурсов	31
Петровский И. И., Свито И. Л., Нехайчик Е. В. Чувствительность фильтров с элементами высшего порядка	33
Сидоров Д., Марков А. В. Способ упреждающего токоограничения электромашин	35
Стешенко П. П., Журавлев В. И. Анализ конструкции и параметров системы межсоединений печатной платы	37
Хаджинов М. К., Павлова А. В., Стасевич Н. А. Сравнение эффективности регуляторов электропривода грузовой тележки при подавлении колебаний подвеса	39

Varchenya I. V., Savenko A. G. Intelligent Storage and Accounting System	41
Секция «Интеллектуальные информационные технологии»	43
Ахапкина А. М., Федоренко В. А. Исследование применения машинного обучения в мобильных приложениях: от проектирования до реализации	45
Гердий Е. В., Зорин Д. О., Жмырко А. В., Банцевич К. А., Орлов М. К. Структура интеллектуальной диалоговой системы для лиц с ограниченными возможностями	47
Ерофеев И. А. Использование интеллектуальных решателей задач в системе управления железнодорожной станцией	49
Зотов Н. В. Количественные показатели эффективности операций над общей семантической памятью интеллектуальных систем	51
Зотов Н. В. Модель управления процессами в общей семантической памяти интеллектуальных систем	53
Ивашенко В. П. Информационные характеристики онтологических структур	55
Ивашенко В. П. Онтологические структуры и параметризованные многозначные логики	57
Казаченко Е. А. Сравнительный анализ различных подходов к проектированию пользовательских интерфейсов	59
Крищеневич В. А., Сальников Д. А., Загорский А. Г., Захарьев В. А. Подходы к анализу биометрических данных в интеллектуальных медицинских системах . .	61
Курочка К. С., Башаримов Ю. С. Сравнительный анализ архитектур RNN для решения задач классификации при обработке видео	63
Мельников Д. А., Петрова А. А. Исследование возможности применения нейронных сетей для оптимизации процессов в системах управления проектами	65
Михальков М. Д. Модели и методы построения интеллектуальных систем поддержки принятия решений . .	67
Никифоров С. А., Садовский М. Е., Гойло А. А. Модуль обработки естественно-языковых сообщений интеллектуальных систем	69

Павлюченко К. А., Киселев А. И. Модели графовых баз данных для анализа предпочтений пользователей	71
Радченко А. Д. Использование искусственного интеллекта в психотерапии	73
Сальников Д. А., Загорский А. Г., Криценович В. А., Кондратьева И. Д., Корсакова С. К. Модели и средства проектирования интеллектуальных систем персонального медицинского обслуживания	75
Смык М. Н., Живицкий И. А. Мобильное устройство для диагностики дисфункции опорно-двигательного аппарата человека	77
Соколов А. П. Оптимизация вычислительных задач машинного обучения на CPU	79
Хирьянов И. Д. Анализ алгоритмов поиска пути в навигационных сетках	81
Хорошавин В. Д., Насевич П. Е., Садовский М. Е. Редактор текстов структурированного представления знаний интеллектуальных систем . .	83
Черепанова Т. В., Корняков К. В. Анатомия искусственного интеллекта	85
Zhidovich A., Lubenko A., Vojteshenko I. About Enhanced Access Control Using FIDO2 Authentication and Attributes	88
Секция «Проектирование встраиваемых систем»	91
Бибило П. Н., Кардаш С. Н. Схемная реализация в заказных СБИС разреженных систем дизъюнктивных нормальных форм булевых функций	93
Бибило П. Н., Кириенко Н. А., Романов В. И. Экспериментальное исследование программ минимизации, выполняемых при синтезе комбинационных схем	95
Бранцевич П.Ю., Лапицкая Н.В., Леванцевич В.А. Регистратор вибрационных сигналов на основе встраиваемой микроконтроллерной системы	97
Габер П. Н., Диваков Н. В. Подходы к построению доверенной среды исполнения в современных вычислительных платформах	99

Кишко Д. В., Алехин А. П. Анализ применения программных эмуляторов в маршруте проектирования СнК	101
Лекомцев А. А., Сухарев А. Д. Инструмент для оценки качества покрытия инструкций компиляторами для архитектуры RISC-V	103
Логинова И. П. Диаграмма Ганта как средство визуализации результатов параллельных программ	105
Можейко Д. О., Иванюк А. А. Генерация случайных чисел на основе дребезга контактов	107
Осипов А. С., Вашкевич М. И. Устройство детектирования и декодирования DTMF-сигналов	109
Силантьев А. М., Примаков Е. В., Переверзев А. Л., Осадчий И. С. Система на кристалле для аппаратно-программной обработки частотно-модулируемых сигналов	111
Черемисинов Д. И., Черемисинова Л. Д. Распознавание логических структур на основе проходных транзисторов в КМОП схеме	113
Черемисинова Л. Д., Черемисинов Д. И. Трансформация иерархического описания КМОП схемы в плоское представление на языке SPICE	115
Шамына А. Ю., Кайкы М. Н., Иванюк А. А. Программно-аппаратный комплекс для быстрого прототипирования и исследования криптографических примитивов на базе ПЛИС	117
Burko L., Kaiky M., Ivaniuk A. Health Tests Hardware Implementation for Entropy Sources	119
Burko L., Kaiky M., Tushinskaya E. Software for Registers Structure HDL-Description Generation with AMBA Protocols Support	121
Kaiky M., Ivaniuk A. Random Number Generation on Reconfigurable Ring Oscillator	123
Kaiky M., Shamyna A., Ivaniuk A. Fast Prototyping of Reconfigurable True Random Number Generation IP-Core	125
Petrovsky D., Ivaniuk A. RISC-V hardware modification for M-sequences generation	127

Секция «Математическое моделирование и компьютерная графика»	129
Баштык П. А., Кукин Д. П., Шабанович Р. А. Алгоритм управления траекторией движения игрового персонажа	131
Горбатов Г. В., Боцьева Э. И. Особенности применения геометрического моделирования, на базе геометрических нод, при решении задач построения трехмерных моделей	133
Гуревич О. В., Кукин Д. П., Новицкая Л. И., Хмыз Д. Д. Рейтинговые системы и их влияние на подбор целевой аудитории при разработке компьютерных игр	135
Журавлёв В. И., Наумович Н. М., Стежко И. К. Моделирование тепловых градиентов в СВЧ плате макета автомобильного радара	139
Журавлёв В. И., Петровский И. И., Свито И. Л. Моделирование взаимного разогрева компонентов силовых преобразователей	141
Кийко В.Н., Ахапкина А. М., Колбун В. С. Моделирование резино-металлических виброизоляторов для синтезаторов частот	143
Кобяк И. П. Интеграл вероятностей ошибки при идентификации выборки числом ВП	145
Коршикова Д. В., Кукин Д. П., Купчина Е. В. Алгоритмы масштабирования графических элементов	147
Коршикова Д. В., Кукин Д. П., Купчина Е. В. Алгоритмы восстановления изображений	149
Курочка К. С., Карась О. В. Алгоритм построения трёхмерной модели черепа человека на основе анализа КТ-изображений	151
Мельникова В. В., Давыдов М. В. Физическое и компьютерное моделирование транскраниальной магнитной стимуляции	153
Новиков П. А. Алгоритмы вычисления коллизий в виртуальном пространстве	155
Панасик А. А., Шилин Л. Ю., Хмыз Д. Д. Способы обеспечения безопасности в виртуальных частных сетях	157
Рязанцев Д. Д., Жилияк Н. А. Алгоритм поиска пути в игровом приложении. Алгоритм Дейкстры	159

Рязанцев Н. Д., Жиляк Н. А. Алгоритм распознавания движений человека	161
Устинович К. А., Лабоцкий В. В. Алгоритмы решения задачи управления складскими запасами	163
Цегельник В. В. Аналитические свойства решений семейства трехмерных динамических пятиэлементных систем с одной квадратичной нелинейностью и четырьмя линейными элементами	165
Zhelenok D. A., Savenko A. G. Automation of Layout and Preparation for Printing Graphic Materials for Board Games	167
Секция «Автоматизированные системы обработки информации»	169
Адамов Н. А., Шилин Л. Ю., Навроцкий А. А. Сверточная нейронная сеть в задачах обработки изображений	171
Акулич В. Н., Шилин Л. Ю., Хмыз Д. Д. Алгоритм Дейкстры при проектировании автоматизированных систем	173
Барановский О. К. Снижение рисков утечки персональных данных путем их обезличивания и анонимизации в системах электронного обучения	175
Болтак С. В., Русакович Е. С. Анализ уязвимостей потоковых криптосистем на базе M-последовательностей	177
Войтович А. С., Ломако А. В. Метапрограммирование в C# с использованием деревьев выражений	179
Высоцкий М. В., Жиляк Н. А., Шабанович Р. А. Скоринговая модель оценки кредитного риска	181
Герман Ю. О., Герман О. В., Мигалевич С. А. Синтез решения на основе логического вывода	183
Ковалевич В. В. Методика расчета структуры и параметров обнаружителя ЛЧМ-сигналов с неизвестными параметрами	185
Лаппо А. И., Ярмолик В. И., Муравьев А. С. Моделирование движения лазера по заданной траектории для процессов формирования переходных отверстий	187
Литвинович А. И., Горячко А. Д. Оценка эффективности внедрения нового поколения образовательных стандартов	189

Микулич Г. В., Жук Е. Е. Статистическая классификация стационарных временных рядов на основе разложения Вольда	191
Пирштук Д. И. Энергоэффективный алгоритм автоматической замены фона для видеоконференций	193
Семижон Е. А. Система поддержки принятия решений (СППР) как инструмент принятия эффективных управленческих решений на предприятиях	195
Скиба И. Г., Нестеренков С. Н., Внук О. М. Нейронные сети для прогнозирования результатов обучения	197
Струкова А. А., магистрант, Шилин Л. Ю., Хмыз Д. Д. Применение метода системного анализа при проектировании автоматизированных систем	199
Судаков Б. Д., Герман О. В. Методы восстановления неполных данных для повышения точности анализа	201
Таранчук В. В. Вопросы интеграции средств LMS Moodle и системы Wolfram Mathematica при создании и развитии интеллектуального контента	203
Хмельницкий В. В. Использование градиентного бустинга над решающими деревьями для прогнозирования временных рядов	205
Шиманский Н. А., Баглов А. В., Хорошко Л. С. Автоматизация обработки результатов исследования структуры материалов	207
Ярмолик В. И., Боброва Т. С., Павловская Е. Р. Производственная практика — как этап в формировании специалистов в области информационных технологий	209
Ярмолик В. И., Боброва Т. С., Ючков А. К. Новый алгоритм выбора лидера при автономном управлении роем БПЛА	211
Yuan H., Gourinovitch A. B. Image Enhancement Algorithms	213
Mukha V. S. Block Methods of Multidimensional-Matrix Data Analysis	215
Umaralizoda R. Sh., Kosimov A. A., Kamolidinov J. B. Features and Characteristics of Artificial Intelligence Systems	219

Savenko A. G., Khotak S. A.	
Web system for automation and digitalization of the processes of organization and support of graduate work design at the graduating department	223
Tang Yi, Zhao Di, Gourinovitch A.B.	
The Redundant Functions Reducing Algorithm	225
Zhao Di,Tang Yi, Gourinovitch A.B.	
Dilated convolution and spatial pyramid fusion in the image segmentation problem	227

Научное издание

**Информационные технологии и системы 2023
(ИТС 2023)**

Материалы международной научной конференции
(БГУИР, Минск, Беларусь, 22 ноября 2023)

**Information Technologies and Systems 2023
(ITS 2023)**

Proceeding of The International Conference
(BSUIR, Minsk, Belarus, 22nd November 2023)

В авторской редакции
Ответственный за выпуск *М. В. Ковалёв*
Компьютерная верстка *Н. В. Гракова*
Дизайн обложки *А. В. Жмырко*

Подписано в печать 16.11.2023. Формат 60x84 1/8. Бумага офсетная. Гарнитура «BookAntiqua».
Отпечатано на ризографе. Усл. печ. л. 41,15. Уч.-изд. л. 45,3. Тираж 35 экз. Заказ 391.

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.