



НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ

САНКТ-ПЕТЕРБУРГСКОГО ГОСУДАРСТВЕННОГО
ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Информатика. Телекоммуникации.
Управление

Том 12, № 4
2019

Санкт-Петербургский политехнический
университет Петра Великого

2019

НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ САНКТ-ПЕТЕРБУРГСКОГО
ГОСУДАРСТВЕННОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА
ИНФОРМАТИКА. ТЕЛЕКОММУНИКАЦИИ. УПРАВЛЕНИЕ

РЕДАКЦИОННЫЙ СОВЕТ ЖУРНАЛА

Председатель

Юсупов Р.М., чл.-кор. РАН;

Редакционный совет:

Абрамов С.М., чл.-кор. РАН;

Арсеньев Д.Г., д-р техн. наук, профессор;

Воеводин В.В., чл.-кор. РАН;

Заборовский В.С., д-р техн. наук, профессор;

Козлов В.Н., д-р техн. наук, профессор;

Фотиади А.Э., д-р физ.-мат. наук, профессор;

Черноруцкий И.Г., д-р техн. наук, профессор.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Главный редактор

Коротков А.С., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Редакционная коллегия:

Ицыксон В.М., канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Prof. Dr. Philippe Ferrari, Head of the RF and Millimeter-Wave Lab IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble Alpes University, France;

Prof. Dr. Wolfgang Krautschneider, Head of Nanoelectronics Institute, Hamburg University of Technology, Germany;

Кучерявый Е.А., канд. техн. наук, профессор, Tampere University of Technology, Finland.

Prof. Dr. Fa-Long Luo, Affiliate Full Professor University of Washington, USA, Chief Scientist Micron Technology, Inc., Milpitas, USA, Chairman IEEE SPS Industry DSP Technology Standing Committee;

Макаров С.Б., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Prof. Dr. Emil Novakov, IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble, France;

Прокопенко Н.Н., д-р техн. наук, профессор, Донской государственный технический университет, г. Ростов-на-Дону, Россия;

Путря М.Г., д-р техн. наук, профессор, Национальный исследовательский университет «Московский институт электронной техники», Москва, Россия;

Пышкин Е.В., канд. техн. наук, доцент, School of Computer Science and Engineering, University of Aizu, Japan;

Трифонов П.В., д-р техн. наук, доцент, Национальный исследовательский университет ИТМО, Санкт-Петербург, Россия;

Устинов С.М., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Уткин Л.В., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Цикин И.А., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Шкодырев В.П., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия.

Журнал с 2002 года входит в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Сведения о публикациях представлены в Реферативном журнале ВИНТИ РАН, в международной справочной системе «Ulrich's Periodical Directory», в базах данных Российский индекс научного цитирования (РИНЦ), Google Scholar, EBSCO, Math-Net.Ru, ProQuest, Index Copernicus

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации ПИ № ФС77-51457 от 19.10.2012 г.

При перепечатке материалов ссылка на журнал обязательна.

Точка зрения редакции может не совпадать с мнением авторов статей.

Адрес редакции: Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.
Тел. редакции (812) 552-62-16.

Санкт-Петербургский политехнический
университет Петра Великого, 2019



ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY JOURNAL

Computer Science.
Telecommunications and Control Systems

Vol. 12, No. 4
2019

ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY JOURNAL COMPUTER SCIENCE. TELECOMMUNICATIONS AND CONTROL SYSTEMS

EDITORIAL COUNCIL

Head of the editorial council

Prof. Dr. *Rafael M. Yusupov* (corresponding member of the Russian Academy of Sciences)

Members:

Prof. Dr. *Sergey M. Abramov* (corresponding member of the Russian Academy of Sciences),

Prof. Dr. *Dmitry G. Arseniev*,

Prof. Dr. *Vladimir V. Voevodin* (corresponding member of the Russian Academy of Sciences),

Prof. Dr. *Vladimir S. Zaborovsky*,

Prof. Dr. *Vladimir N. Kozlov*,

Prof. Dr. *Alexandr E. Fotiadi*,

Prof. Dr. *Igor G. Chernorutsky*.

EDITORIAL BOARD

Editor-in-chief

Prof. Dr. *Alexander S. Korotkov*, Peter the Great St. Petersburg Polytechnic University, Russia;

Members:

Assoc. Prof. Dr. *Vladimir M. Itsykson*, Peter the Great St. Petersburg Polytechnic University, Russia;

Prof. Dr. *Philippe Ferrari*, Head of the RF and Millimeter-Wave Lab IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble Alpes University, France;

Prof. Dr. *Yevgeni Koucheryavy*, Tampere University of Technology, Finland.

Prof. Dr. *Wolfgang Krautschneider*, Head of Nanoelectronics Institute, Hamburg University of Technology, Germany;

Prof. Dr. *Fa-Long Luo*, Affiliate Full Professor University of Washington, USA, Chief Scientist Micron Technology, Inc., Milpitas, USA, Chairman IEEE SPS Industry DSP Technology Standing Committee;

Prof. Dr. *Sergey B. Makarov*, Peter the Great St. Petersburg Polytechnic University, Russia;

Prof. Dr. *Emil Novakov*, IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble, France;

Prof. Dr. *Nikolay N. Prokopenko*, Don State Technical University, Rostov-on-Don, Russia;

Prof. Dr. *Mikhail G. Putrya*, National Research University of Electronic Technology, Moscow, Russia;

Sen. Assoc. Prof. Dr. *Evgeny Pyshkin*, School of Computer Science and Engineering, University of Aizu, Japan;

Prof. Dr. *Viacheslav P. Shkodyrev*, Peter the Great St. Petersburg Polytechnic University, Russia;

Prof. Dr. *Peter V. Trifonov*, ITMO University, St. Petersburg, Russia;

Prof. Dr. *Igor A. Tsikin*, Professor, Peter the Great St. Petersburg Polytechnic University, Russia;

Prof. Dr. *Sergey M. Ustinov*, Peter the Great St. Petersburg Polytechnic University, Russia;

Prof. Dr. *Lev V. Utkin*, Peter the Great St. Petersburg Polytechnic University, Russia.

The journal is included in the List of Leading Peer-Reviewed Scientific Journals and other editions to publish major findings of PhD theses for the research degrees of Doctor of Sciences and Candidate of Sciences.

The journal is indexed by Ulrich's Periodicals Directory, Google Scholar, EBSCO, ProQuest, Index Copernicus, VINITI RAS Abstract Journal (Referativnyi Zhurnal), VINITI RAS Scientific and Technical Literature Collection, Russian Science Citation Index (RSCI) database © Scientific Electronic Library and Math-Net.ru databases.

The journal is registered with the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR). Certificate ПИ № ФС77-51457 issued Oct. 19, 2012.

No part of this publication may be reproduced without clear reference to the source.

The views of the authors can contradict the views of the Editorial Board.

The address: 195251 Polytekhnicheskaya Str. 29, St. Petersburg, Russia.

Peter the Great St. Petersburg
Polytechnic University, 2019

Содержание

Информационные технологии

Попов С.Г., Пономарева Е.В., Самочадин А.В. Исследование виджета стратегирования в системе интерактивного бизнес-анализа	7
---	---

Телекоммуникационные системы и компьютерные сети

Фатеева А.С. Частотно-независимый способ цифровой калибровки каналов АЦП устройств цифровой обработки сигналов	20
---	----

Программное обеспечение вычислительных, телекоммуникационных и управляющих систем

Левченко А.В. Метод гибридного неоднородного тайлинга для архитектур суперкомпьютеров с многоуровневой иерархией памяти	29
Улитин А.В., Смирнов А.А. Программный комплекс управления материальными ресурсами листового металла с элементами СППР	45

Устройства и системы передачи, приёма и обработки сигналов

Рашич А.В., Горбунов С.В. Итеративный приём многочастотных сигналов с неортогональным частотным уплотнением в каналах с замираниями	58
Боробов А.А., Бородулин Р.Ю. Методика оценки интермодуляционных помех группы радиостанций средней мощности	72
Березняк А.Ф., Коротков А.С. Синтез и реализация монолитных интегральных схем СВЧ-переключателей на основе GaAs pHEMT-технологии	84

Моделирование вычислительных, телекоммуникационных, управляющих и социально-экономических систем

Беляевский К.О. Формирование октодерев по облаку точек при ограничении объёма оперативной памяти	97
Глазунов В.В., Чуватов М.В., Чернышев А.С., Курочкин Л.М. Метод и технология интеграции дискретной и сплошноредной моделей транспортных потоков региона	111

Системный анализ и управление

Родионов Д.В., Люхтер А.Б., Прокошев В.Г. Расчет точек технологической траектории лазерного роботизированного комплекса относительно инструмента и оснастки в САМ	123
Филиповский В.М., Полюга Н.Л., Мухамбедьяров Б.Б. Синтез дискретного регулятора переоборудованием с учётом задержки в вычислительном устройстве	136
Успенский М.Б. Обзор подходов к обнаружению сбоев в системах хранения данных	145

Contents

Information Technologies

Popov S.G., Ponomareva E.V., Samochadin A.V. Researching a strategy widget in the business analysis interactive system	7
---	---

Telecommunication Systems and Computer Networks

Fateeva A.S. Frequency-independent method of digital calibration of ADC channels digital signal processing devices	20
---	----

Software of Computer, Telecommunications and Control Systems

Levchenko A.V. A hybrid non-uniform loop tiling for supercomputers with deep memory hierarchies	29
Ulitin A.V., Smirnov A.A. Software package for managing material resources of sheet metal with elements of DSS	45

Hardware of Computer, Telecommunications and Control Systems

Rashich A.V., Gorbunov S.V. Iterative receiver of SEFDM signals in fading channels	58
Borobov A.A., Borodulin R.Yu. Method for assessing intermodulation interference of a group of medium-power radio stations	72
Berezniak A.F., Korotkov A.S. Design method and manufacturing monolithic microwave integrated circuit switches on GaAs pHEMT	84

System Analysis and Control

Beliaevskii K.O. Generating octrees from the point cloud in conditions of limited random access memory volume	97
Glazunov V.V., Chuvatov M.V., Chernyshev A.S., Kurochkin L.M. The method and technology for integrating discrete and continuous models of regional traffic flows	111

Intellectual Systems and Technologies

Rodionov D.V., Lyukhter A.B., Prokoshev V.G. Calculation of the points of the technological trajectory of the laser robotic complex relative to the tool and tooling in the cam	123
Filipovskii V.M., Polyuga N.L., Mukhamedyarov B.B. Synthesis of the discrete regulator by re-equipment method taking into account the delay in a computer device	136
Uspenskiy M.B. A survey of the approaches to storage systems fault detection	145

DOI: 10.18721/JCSTCS.12401
УДК 004.652.4

ИССЛЕДОВАНИЕ ВИДЖЕТА СТРАТЕГИРОВАНИЯ В СИСТЕМЕ ИНТЕРАКТИВНОГО БИЗНЕС-АНАЛИЗА

С.Г. Попов¹, Е.В. Пономарева², А.В. Самочадин¹

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация;
² «ИБС Экспертиза»,
Санкт-Петербург, Российская Федерация

Рассмотрены подходы к решению задачи интерактивного стратегирования и бизнес-анализа с элементами прогнозирования. Цель исследования — определение возможных вариантов решения задач, возникающих при разработке виджета интерактивного стратегирования и бизнес-анализа с элементами прогнозирования для коммерческих организаций. Результатом является разработанный виджет для интерактивного стратегирования и бизнес-анализа с элементами прогнозирования. Описаны аналоги, приведена технология создания виджета, реализация виджета, а также проведено исследование зависимости времени отображения элементов виджета и виджета целиком от количества стратегических целей и показателей в современных веб-браузерах.

Ключевые слова: архитектура системы бизнес-анализа, виджет, web-клиент, интерактивное стратегирование, интерактивные панели, архитектура приложения.

Ссылка при цитировании: Попов С.Г., Пономарева Е.В., Самочадин А.В. Исследование виджета стратегирования в системе интерактивного бизнес-анализа // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 7–19. DOI: 10.18721/JCSTCS.12401

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

RESEARCHING A STRATEGY WIDGET IN THE BUSINESS ANALYSIS INTERACTIVE SYSTEM

S.G. Popov¹, E.V. Ponomareva², A.V. Samochadin¹

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation;
² IBS Expertise,
St. Petersburg, Russian Federation

The article discusses approaches to solving problems of interactive strategy and business analysis with forecasting elements. The aim of the study is to solutions to problem that arise when developing a widget for interactive strategy and business

analysis with forecasting elements for government and commercial organizations. The result is a widget for interactive strategy and business analysis with forecasting elements. The article discusses the analogues, describes the technology for creating a widget, to realizes the widget, and also studies the dependence of the display time of widget and widget elements entirely on the number of strategic goals and indicators in modern web browsers.

Keywords: business intelligence system architecture, widget, web client, interactive strategy, interactive panels, application architecture.

Citation: Popov S.G., Ponomareva E.V., Samochadin A.V. Researching a strategy widget in the business analysis interactive system. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 7–19. DOI: 10.18721/JCSTCS.12401

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

В условиях современной рыночной экономики стратегирование является важным процессом, позволяющим вычислить основные направления развития предприятия и показатели его развития на планируемый период, определяющие желаемый результат его деятельности. Стратегирование можно описать как совокупность процессов стратегического планирования, прогнозирования и управления [7], которые приводят к формированию эффективного плана действия для достижения определенных целей организации. Стратегическое планирование помогает определить цели организации, выделить для них необходимые ресурсы и эффективно распределить их для получения наибольшей выгоды. Прогнозирование помогает оценить возможные значения показателей в будущем и способствует формированию достижимых целей при разработке стратегии.

В настоящий момент в крупных корпорациях внедряются системы бизнес-аналитики, отвечающие за извлечение, преобразование, загрузку и хранение данных. Для отображения данных эти системы используют графические элементы — виджеты, позволяющие пользователю строить визуализацию данных на основе хранимых в информационной системе значений. Набор из взаимосвязанных виджетов может быть сохранен на аналитической па-

нели для дальнейшего просмотра и анализа. Все виджеты на панелях являются интерактивными, что позволяет пользователю взаимодействовать с данными.

Для разработки концепции, прогнозов и стратегии в системах бизнес-аналитики используются такие методы, как индикативное и стратегическое планирование [6]. Стратегическое планирование помогает компании выделить четкие цели, разработать способы их достижения и определить ресурсы, необходимые для реализации данной цели. Вопросы использования стратегического планирования и прогнозирования рассмотрены в трудах отечественных и зарубежных авторов [1, 5, 8]. Технология помогает руководителям эффективно принимать важные управленческие решения, предоставляя мгновенный доступ к аналитическим панелям с визуальной информацией о текущих целях компании, ключевых показателях эффективности, трендах, зависимостях и других метриках, что способствует быстрому выявлению источника проблем и нахождению способов их решения.

Аналитические панели помогают сравнить конкретные показатели во времени или оценить их относительно других показателей. Аналитические данные при этом являются сгруппированными и агрегированными исходными данными, которые могут иметь различные визуальные представления, такие как таблица, диаграмма,

индикатор или график. Основным инструментом для реализации стратегирования является визуализация данных, т. к. средства визуализации являются неотъемлемым элементом BI-платформ, помогая пользователю увидеть общую картину, скрывающуюся за данными, и способствуя принятию правильных управленческих решений. При этом основным подходом к реализации средств визуализации будут являться интерактивные виджеты.

Программная реализация виджета стратегирования в системах бизнес-аналитики

Современные системы бизнес-аналитики обладают широкими возможностями реализации интерактивного стратегирования, основанными на прогнозах и КПЭ [18], при этом являясь либо системами управления стратегиями предприятия, основанными исключительно на КПЭ или на универсальных приёмах построения интерактивных панелей [15, 16].

KPI Монитор — программная платформа, которая применяется для своевременной оценки производительности работы компании при помощи автоматизации системы управления КПЭ. Данная система разработана российской компанией ПроФитПроект, специализирующейся на разработке и внедрении учетно-аналитических программ для автоматизации бухгалтерского и управленческого учета [9].

Power BI — аналитическая платформа для бизнес-анализа, разработанная Microsoft, объединяющая в себе несколько программных продуктов, таких как Power BI Desktop, Power BI Services, Power BI Mobile. В Power BI нет визуального элемента для стратегирования, однако комбинация элементов для KPI и прогнозирования может помочь в реализации построения стратегии, а для наблюдения за KPI были разработаны средства Power KPI, Power KPI Power KPI Indicator [11]. Power BI является более наглядным элементом при анализе КПЭ, т. к. помимо построения графиков на нем отображается дополнительная информация

о текущей дате, текущем значении показателя и статусе выполнения целевого значения показателя.

Prognoz Platform — платформа бизнес-аналитики, разработанная компанией «Прогноз». Данная платформа поддерживает такие технологии, как настольное приложение, веб-приложение, мобильное приложение. Все данные могут быть представлены в виде аналитических панелей и отчетов. Для формирования стратегических планов, их модификации и оценки достижения целей в Prognoz присутствует инструмент «Карты показателей» [10].

Таким образом, в каждой из представленных систем есть возможность в каком-либо виде визуально представить стратегию на основании информации о КПЭ, построить график по истории значений КПЭ и сохранить полученную стратегию на аналитической панели для дальнейшего анализа. Основные отличия платформ представлены в табл. 1.

Можно сделать вывод, что в KPI Monitor есть хорошие инструменты для построения стратегических карт, неограниченная древовидная вложенность системы показателей, интерактивное раскрытие узлов, позволяющее проводить многомерный аналитический анализ с выяснением причинно-следственных связей. Однако в системе полностью отсутствуют какие-либо методы прогнозирования, которые являются важной составляющей стратегирования.

Удобство Power BI заключается в наличии инструментов для построения сбалансированных систем показателей с определением статуса выполнения и построением графиков для неограниченного количества КПЭ в одном визуальном элементе. Также система предоставляет на выбор несколько методов построения прогноза, однако, при построении стратегии пользователь не может выделить определенные цели и показать зависимость этих целей друг от друга и от различных показателей.

Таблица 1

Сравнительная характеристика виджетов стратегирования современных систем бизнес-аналитики

Table 1

Comparative characteristics of interactive widgets of modern business intelligence systems

	KPI Монитор	Power BI	Prognoz Platform
Доступные способы прогнозирования	—	Линейная регрессия, линейный тренд, скользящее среднее, простое экспоненциальное сглаживание	Автотренд, геометрический тренд, грей-метод, линейный тренд, обратный тренд, скользящее среднее, простое экспоненциальное сглаживание
Период прогноза	—	∞	До 2035
Наличие визуального индикатора выполнения цели	+	+	—
Представление стратегии	Стратегическая карта	Таблица КПЭ	Стратегическая карта, таблица КПЭ
Способ подсчёта выполнения цели	Весы	—	Формула

Prognoz Platform предоставляет возможность построения стратегических карт, определения перспектив, целей, КПЭ и создания правила достижения определенной цели. В нём реализовано два вида представления стратегии: графическая схема и таблица целей и показателей, отображающая фактические и плановые значения показателей, а также их соотношение. При просмотре показателей пользователь может построить графики для различных показателей. В инструменте «Анализ временных рядов» данной программы может быть построен прогноз для определённого показателя, однако сам инструмент «Карта показателей» не предоставляет такой возможности. Также отсутствует возможность добавления визуального индикатора выполнения цели.

Постановка задачи. Цель нашей работы — исследование реализованного средства визуализации интерактивного стратегирования — виджета «Стратегическая карта», которая обеспечивает группировку целей по четырем стратегическим перспективам, вывод информации о выполнении цели, отображение индикатора выполне-

ния цели, отображение планового и фактического значения показателей, визуальное отслеживание изменения показателя во времени, изменение периода отображения показателя во времени, расчёт прогнозируемых значений для выбранных пользователем показателей с применением одного из трёх методов прогнозирования (линейная регрессия, квадратичная регрессия, тройное экспоненциальное сглаживание). Данные о цели, индикаторы и показатели, а также результат прогноза визуально отображаются на панели.

Технология создания виджета интерактивного стратегирования

В настоящее время стратегирование основано на применении сбалансированной системы показателей (ССП), позволяющей оценить эффективность компании при помощи оптимально подобранных и сбалансированных показателей, которые комплексно отражают ее текущее состояние [2]. Правильно разработанная СПП не только предоставляет руководителям своевременный доступ к информации о состоянии компании на текущий момент, но

эффективности [17]. Использование сочетания виджетов прогнозирования и оценок КПЭ позволяет построить набор визуальных моделей для поддержки принятия решений [20].

Ориентируясь на реализацию универсальной системы бизнес-аналитики, к дальнейшему рассмотрению предложена информационная панель стратегирования, состоящая из четырех стратегических направлений и двух показателей, отражающих предпочтения руководства и сложившиеся бизнес-практики конкретного предприятия. Диаграмма вариантов использования в процессе создания и использования виджета для акторов «Аналитик» и «Пользователь» представлена на рис. 1.

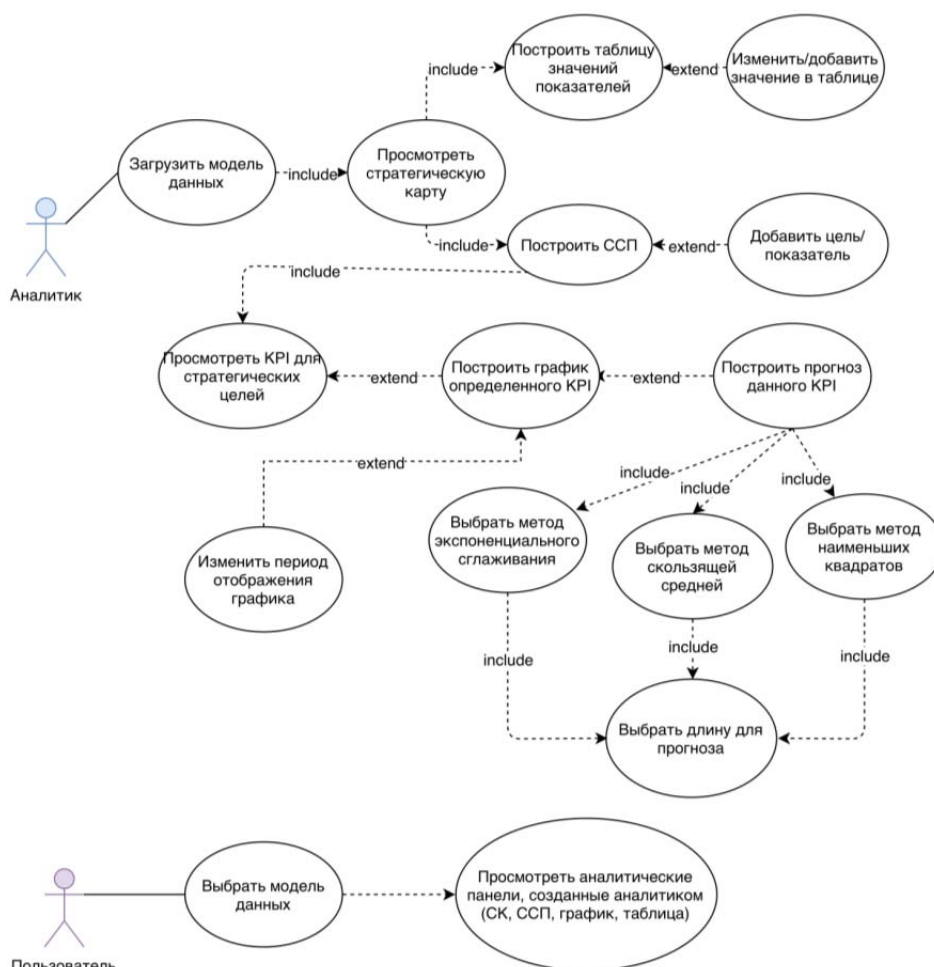


Fig. 1. Strategy widget usage diagram

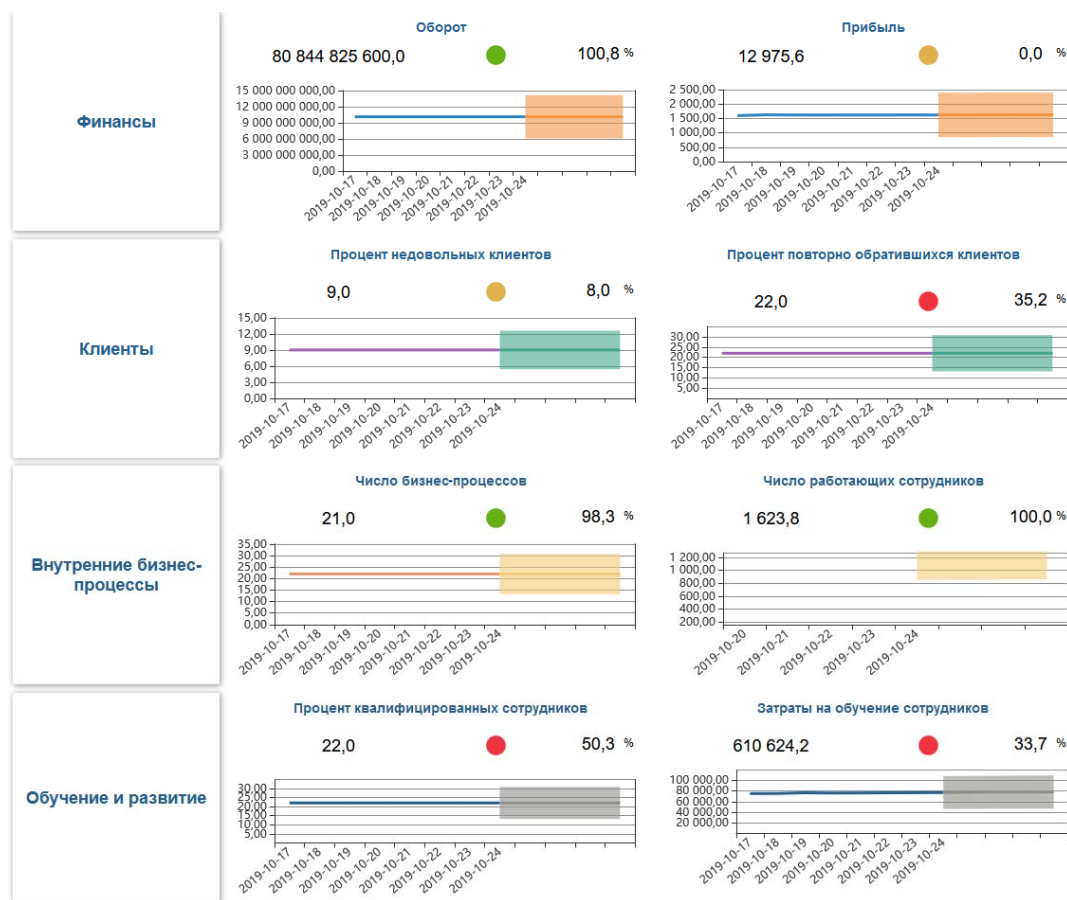


Рис. 2. Пример отображения виджета стратегической карты

Fig. 2. An example of displaying a strategic map widget

Для каждого сочетания стратегического направления и показателя реализованы визуальные модели прогнозирования (прогноз временного ряда), требуемые и достигаемые на конкретный момент КПЭ. Пример интерфейса сформированного интерактивного виджета стратегирования приведён на рис. 2.

Виджет прогнозирования обеспечивает представление временного ряда стратегического направления в разрезе каждого показателя, а значения КПЭ показывают заданные показатели на текущий момент, окончание периода построения стратегии, и качественную оценку состояния показателя.

Для реализации средства визуализации разработана функциональная диаграмма, представленная на рис. 3. Диаграмма позволяет отобразить связи между функциями разрабатываемого виджета. Каждый блок диаграммы является функцией, у которой

есть входные и выходные данные и управляющая информация. Иерархическое представление блоков на диаграмме показывает последовательность их вызова в программе и влияние друг на друга.

Место виджета прогнозирования и интерактивного стратегирования в системе бизнес-аналитики приведено на структурной схеме системы интерактивного стратегирования на рис. 4.

Виджет получает агрегированные данные от подсистемы хранения данных и аналитик, основываясь на бизнес-планах компании строит сбалансированную систему показателей, которая затем отображается для пользователя. Процесс взаимодействия аналитика и пользователя с виджетом состоит из шести этапов.

1. Загрузка данных в виджет для построения стратегии.

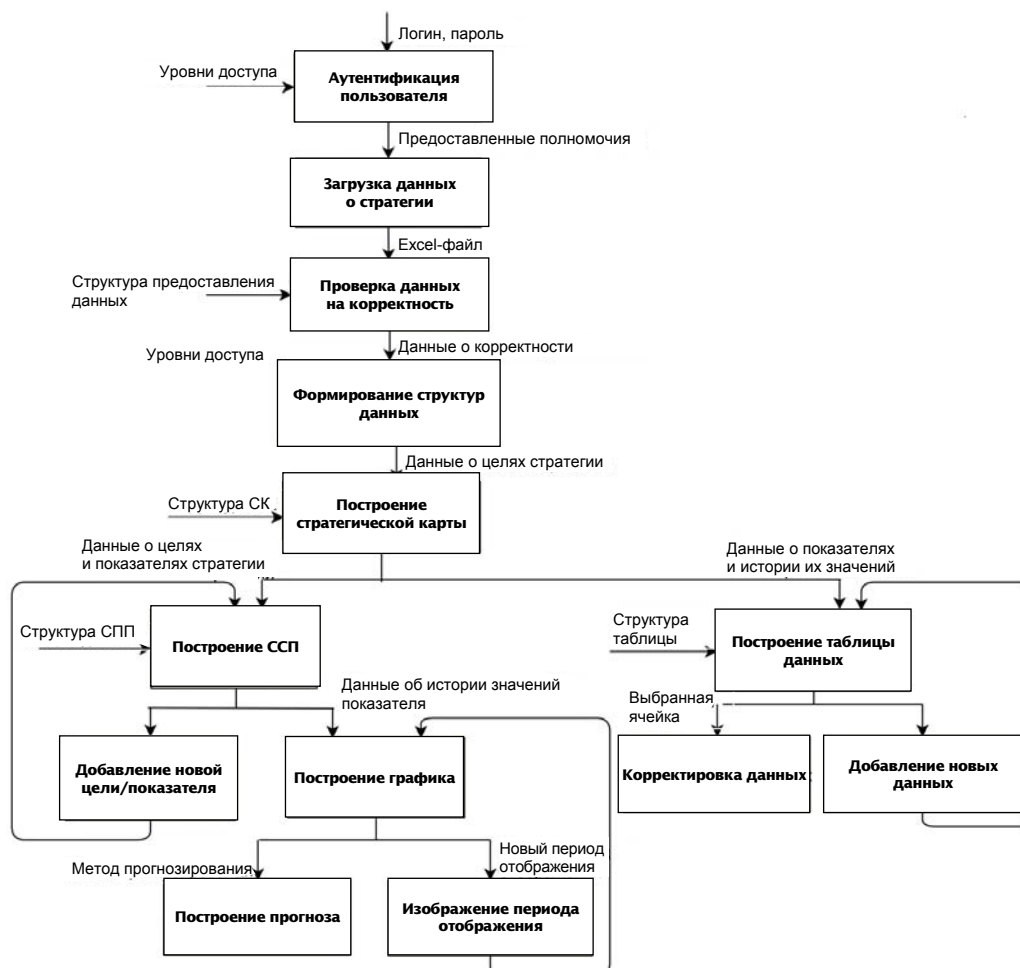


Рис. 3. Функциональная схема информационной панели стратегирования

Fig. 3. Block diagram of the strategic information panel



Рис. 4. Структурная схема системы интерактивного стратегирования

Fig. 4. Block diagram of an interactive strategy system

2. Сохранение данных и выполнение вычислений.

3. Добавление или модификация данных.

4. Переход от редактирования стратегии к просмотру.

5. Запрос на отображение виджета.

6. Формирование содержимого страницы и вывод виджета с запрошенными данными в браузере на стороне клиента.

Реализация виджета обеспечивает его функционирование в двух режимах: редактирования и просмотра. В режиме редактирования для аналитика доступны функции выбора данных, редактирования показателей и подстройки методов прогноза, тогда как для пользователя доступны только функции просмотра показателей, текущих значений и значений прогноза.

Реализация виджета интерактивного стратегирования

В качестве архитектуры для разработки клиентского приложения была выбрана Model-View-Controller (MVC), позволяющая разделить функции виджета на три блока, взаимосвязанных между собой: «Модель», «Представление» и «Контроллер».

Модель отвечает за хранение и обработку данных. В блоке «Модель» хранятся две структуры объектов, представленных как ассоциативные массивы, необходимые для корректного отображения стратегической карты. На первом уровне структуры находится список из названий ключевых целей стратегии. Для каждой цели в формате ключ-значение хранится ее описание, тип перспективы, информация о минимальном и максимальном плановом значении, текущем проценте выполнения и подцелях, от которых она зависит. Подцели могут являться как самостоятельными целями, так и показателями, что определяется их типом. Каждая подцель описывается именем, уникальным идентификатором, плановым значением на ближайший месяц, весом и необходимой тенденцией из-

менения, к которым относятся рост или падение. Сумма всех весов подцелей для цели равна единице, что позволяет определить процент ее выполнения. На втором уровне структура отвечает за корректное построение графиков. Для каждого показателя хранится информация о его значении за каждый месяц в определенные года. В модели реализован метод добавления новой цели или показателя и преобразования данных из внутреннего формата хранения в формат модели.

Блок «Представление» отвечает за отображение данных из модели посредством взаимодействия с объектной моделью документа (DOM). При этом если данные модели изменяются, представление также должно обновляться по мере необходимости. Для реализации виджета стратегической карты использованы базовые элементы HTML, к которым относятся, например, текстовые блоки, формы и списки, стилизованные с помощью каскадных таблиц стилей (CSS), и управляемые посредством языка программирования JavaScript. Окончательное форматирование отображения стратегической панели на стороне клиента осуществляется при помощи библиотеки Echarts, поскольку она обладает высокой производительностью по сравнению с аналогами.

Блок «Контроллер» является связующим звеном между моделью и представлением. После того как произошло определенное событие, контроллеру необходимо связаться с блоком модели для получения или модификации данных, после чего предоставить их блоку представления для отображения полученного результата.

Исследование времени отображения данных виджета стратегической карты

Основным параметром качества функционирования виджета интерактивного стратегирования является его интерактивность при сохранении кросс-браузерности [13], где под интерактивностью подразумевается малое время реакции системы на действия пользователя в современных вер-

сиях распространенных браузеров. Цель проводимых экспериментов – определение числа стратегических перспектив и показателей, при котором сохраняется интерактивный ответ системы, составляющий 150–200 мс [12]. Исследование времени проводилось для операции загрузки данных, построения графиков прогнозов и выполнения виджета в целом. Технические характеристики компьютера, на котором проводилось тестирование, приведены в табл. 2.

Для исследования времени загрузки данных использовались тестовые файлы с подготовленными данными для выбранного числа показателей. Загрузка каждого файла выполнялась отдельно, после чего оценивалось среднее время загрузки. Результаты эксперимента представлены в табл. 3. Видно, что время загрузки данных для генерации отображения практически не зависит от числа целей. Все браузеры загрузили файлы на 183 цели и 512 значений показателей для каждой цели менее чем за 1 мс.

Таблица 2

Технические характеристики рабочей станции пользователя

Table 2

Technical characteristics of the workstation

Название параметра	Значение
Тип процессора, тактовая частота, ГГц	Intel Core i5, 1.6
Объём оперативной памяти, Гб	8
Тип видеокарты, оперативной памяти, Гб	NVIDIA GeForce MX150, 2
Операционная система	iOS, Windows
Производитель браузера, визуализатор, версия	Google Chrom, Blink, 74.0.3729.169; Safari, WebKit, 60.0.3255.118; Firefox Quantum, Gecko, 67.0; Microsoft Edge, EdgeHTML, 42.17134.1.0.

Таблица 3

Время загрузки данных в виджет в различных браузерах

Table 3

Download time of data to the widget in various browsers

Число целей	Время обработки в браузере, мс			
	Google Chrome	Firefox	ME	Safari
3	0,8	0,3	0,3	0,2
5	0,8	0,4	0,5	0,3
8	0,8	0,6	0,3	0,2
14	0,9	0,4	0,4	0,5
24	0,8	0,3	0,4	0,4
46	0,8	0,5	0,8	0,4
89	0,8	0,4	0,7	0,5
183	0,9	0,4	0,6	0,4

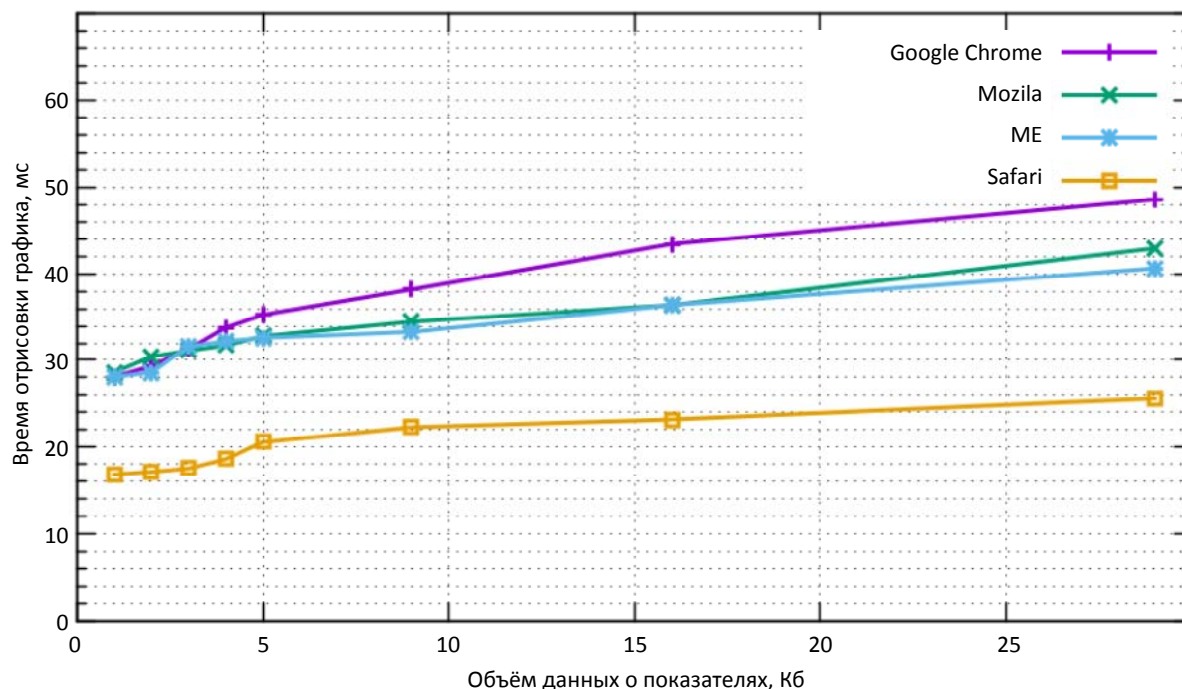


Рис. 5. Зависимость времени отображения графика прогноза от числа данных показателей в различных браузерах

Fig. 5. The dependence of the display time of the forecast graph on the number of indicators in different browsers

Таким образом, несмотря на то что Microsoft Edge и Google Chrome показали на тестах наилучшие результаты, можно сказать, что все браузеры обеспечивают интерактивное взаимодействие пользователя при работе с основными функциональными возможностями виджета без каких-либо видимых задержек.

Исследование времени отображения графиков прогнозов в браузерах выполнено для тестовых наборов данных в диапазоне от 4 до 512 Кбайтов. Количество показателей в одной серии увеличивалось от 10 до 30, чтобы повысить точность получаемых значений, т. к. разброс между значениями оказался достаточно мал. График зависимости времени отображения графиков прогноза приведён на рис. 5.

На основании данных можно сказать, что Safari снова оказался лидером по скорости, тогда как остальные браузеры справились с задачей примерно одинаково. В результате исследования выяснилось, что

количество значений показателя не накладывает ограничений на интерактивность построения графика.

Для определения максимального объёма данных, при котором разработанный виджет остается интерактивным для пользователя, было проведено тестирование с использованием заранее подготовленных наборов данных, путём создания различных по объёму файлов формата CSV. Для отображения данных в виджете были созданы файлы, количество целей в которых изменялось от 4 до 512, причем каждая цель зависела от последующей цели и от одного ключевого показателя эффективности. Количество выгружаемых ячеек при этом составило от 56 до 7674. Размер файлов изменился с 2 до 154 Кб. В каждой серии проводилось 12 экспериментов и вычислялось среднее значение времени отображения. Графики зависимости времени отображения в различных браузерах приведены на рис. 6.

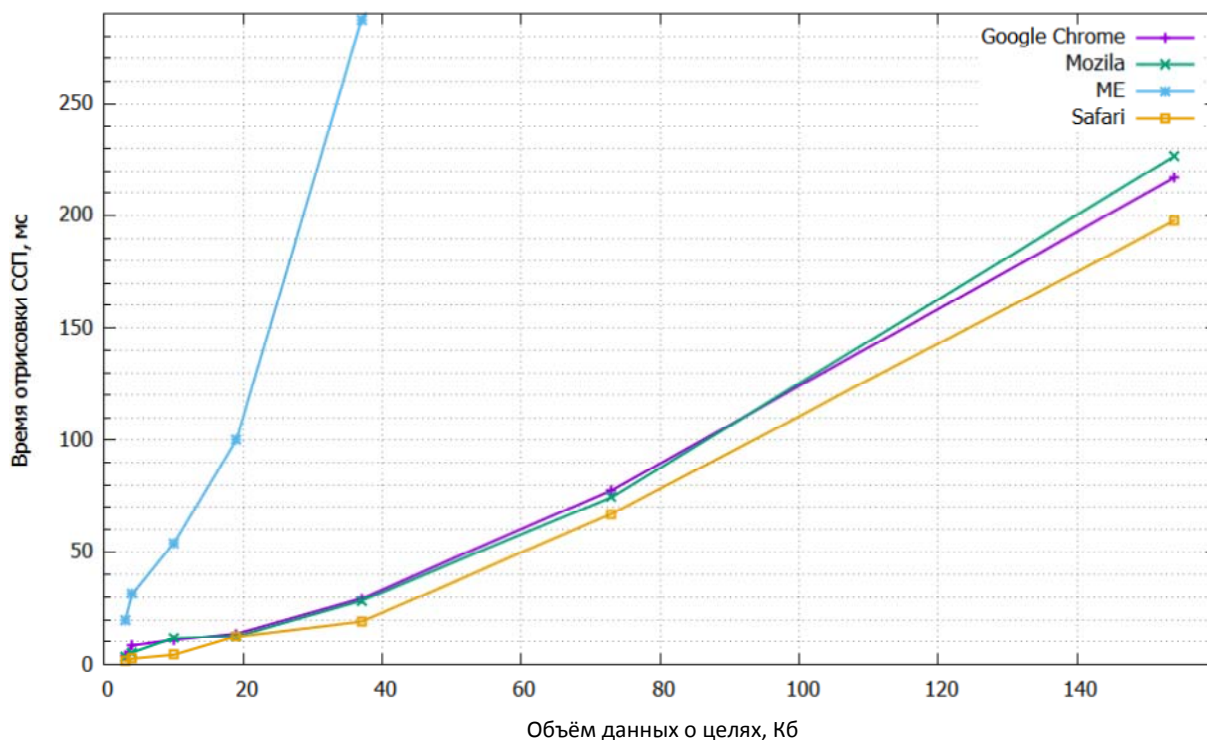


Рис. 6. Зависимость времени отображения панели стратегирования от числа перспектив в различных браузерах

Fig. 6. Dependence the display time of the strategic panel on the number of perspectives in various browsers

На основании данных графика можно сделать вывод, что с отображением панели все браузеры справились. Самые быстрые результаты показал браузер Safari. Более медленными оказались браузеры Google Chrome и Mozilla, однако все три браузера позволяют построить виджет с 256 связанными между собой целями без каких-либо видимых задержек для пользователя. Microsoft Edge показал не такие хорошие результаты, однако он сохранил интерактивность при отображении 64 целей, что достаточно для описания стратегических целей компании.

Заключение

В статье рассмотрены технология взаимодействия, архитектурные решения и изложены результаты исследования виджета интерактивного стратегирования в системах бизнес-анализа. В процессе реализации собственного решения разработана

модель взаимодействия аналитика и пользователя с виджетом, сформулирован сценарий взаимодействия, продемонстрированы макеты интерфейса и приведены интерактивные элементы виджета. Реализована архитектура виджета, состоящая из трёх взаимосвязанных блоков: «Модель», «Представление» и «Контроллер».

Для проверки границ применимости и работоспособности виджет протестирован в четырёх современных браузерах, для которых были проведены эксперименты по исследованию времени выполнения ключевых процессов: загрузки данных, отображения графика прогноза и виджета в целом. В ходе тестирования выявлено, что максимальное время ответа на воздействие пользователя не превышало 250 мс, что позволяет констатировать сохранение свойства интерактивности панели при широком диапазоне числа взаимозависимых целей. Рассмотренные браузеры

предоставляют комфортное время взаимодействия пользователя с виджетом, без видимых задержек, что показывает эффективность предложенных архитектурных решений. В дальнейшем набор возможностей виджета может быть расширен использованием дополнительных элементов отображения данных, к которым относятся круговые и столбчатые диаграм-

мы, таблицы, что позволит аналитикам и пользователям принимать управленческие решения для реализации стратегии развития компании.

Работа подготовлена в ходе реализации комплексного проекта в рамках Постановления Правительства РФ от 09.04.2010 № 218 при финансовой поддержке Минобрнауки РФ. Договор № 03.G25.31.0259 от 28.04.2017.

СПИСОК ЛИТЕРАТУРЫ

1. Кузык Б.И., Яковец Ю.В., Кушлин В.И. Прогнозирование, стратегическое планирование и национальное программирование. М.: Экономика, 2011. 604 с.
2. Гуськова Н.Д., Краковская И.И., Вдовин С.М. Место стратегирования в понятийно-категориальной системе прогнозирования // Национальные интересы: приоритеты и безопасность. 2013. № 20.
3. Зук К., Аллен Д. Стратегии роста компании в эпоху нестабильности. М.: Вильямс, 2007. 189 с.
4. Роберт К., Дейвид Н. Сбалансированная система показателей. От стратегии к действию. М.: Олимп-Бизнес, 2017. 320 с.
5. Виханский О.С. Стратегическое управление. М.: Экономистъ, 2008. 296 с.
6. Зельднер А.Г. Место стратегирования в понятийно-категориальной системе прогнозирования // Экономические науки. 2012. № 93. С. 7–15.
7. Цыбатов В.А. Стратегирование регионального развития: методы, модели, информационные технологии // Региональная экономика: теория и практика. 2015. Vol. 13. № 27. С. 36–52.
8. Грант Р. Современный стратегический анализ. СПб.: Питер, 2018. 554 с.
9. Сбалансированная система показателей KPI monitor // URL: <http://www.kpi-monitor.ru/solutions/balanced-scorecard/> (Дата обращения: 23.05.2019).
10. Карты показателей prognos platform // URL: http://help.prognos.com/8.0/ru/mergedProjects/UiScoreCard/uiscorecard_title.htm/ (Дата обращения: 23.05.2019).
11. Koehler G. Balanced scorecards in power BI. 2018. January // URL: <https://powerbi.microsoft.com/ru-ru/blog/balanced-scorecards/> (Дата обращения: 23.05.2019).
12. Nah F. A study on tolerable waiting time: how long are web users willing to wait? // Behaviour and Information Technology. 2004. Vol. 23. No. 3. Pp. 153–163.
13. Nielsen J. Website response times. 2010. June. // URL: <https://www.nngroup.com/articles/website-response-times/> (Дата обращения: 23.05.2019).
14. Rigby D., Bilodeau B. Management tools and trends 2015. June // URL: http://www2.bain.com/management_tools/BainTopTenTools/2017/default.asp/ (Дата обращения: 23.05.2019).
15. Khedkar S. Atul, Shinde S.K. Customer review analytics for business intelligence // 2018 IEEE International Conference on Computational Intelligence and Computing Research. Madurai, India, 2018. Pp. 1–5.
16. Kamimori S., Ogata S., Kaijiri K. Automatic method of generating a web prototype employing live interactive widget to validate functional usability requirements // 2015 3rd International Conference on Applied Computing and Information Technology / 2nd International Conference on Computational Science and Intelligence. Okayama, 2015. Pp. 8–13.
17. Гаврилова Т.А. и др. Об одном методе классификации визуальных моделей // Бизнес-информатика. 2013. № 4 (26). С. 21–34.
18. Шеботинов А.А., Колычев В.Д. Визуализация стратегических KPI организации с использованием инструментальных средств // Скиф. Вопросы студенческой науки. 2018. № 4 (20).
19. Колычев В.Д., Румянцев В.П. Система визуальных моделей управления проектами // Научная визуализация. Электронный журнал. 2014. Т. 6. № 3. С. 14–54.
20. Цветков В.Я. Визуальное моделирование в системах поддержки принятия решений // Международный журнал прикладных и фундаментальных исследований. 2016. № 10-1. С. 13–17.

Статья поступила в редакцию 08.11.2019.

REFERENCES

1. **Kuzyk B.I., Yakovets Yu.V., Kushlin V.I.** *Forecasting, strategic planning and national programming*. Moscow: Ekonomika Publ., 2011. 604 p. (rus)
2. **Guskova N.D., Krakovskaya I.I., Vdovin S.M.** *The strategy's position in the conceptual-categorical forecasting system*. 2013, P. 20. (rus)
3. **Zuk K., Allen D.** *Company growth strategies in an era of instability*. Moscow: Vilyams Publ., 2007. 189 p. (rus)
4. **Robert K., Deyvid N.** *Balanced scorecard. From strategy to action*. Moscow: Olimp-Biznes Publ., 2017. 320 p. (rus)
5. **Vikhanskiy O.S.** *Strategic management*. Moscow: Ekonomist Publ., 2008. 296 p. (rus)
6. **Zeldner A.G.** The place of strategy in the conceptual-categorical forecasting system. *Economics*, 2012, No. 93, Pp. 7–15. (rus)
7. **Tsybatov V.A.** Strategic planning of regional development: methods, models, information technology. *Regional Economics: Theory and Practice*, 2015, Vol. 13, No. 27, Pp. 36–52. (rus)
8. **Grant R.** *Modern strategic analysis*. St. Petersburg: Piter Publ., 2018. 554 p. (rus)
9. Sbalansirovannaya sistema pokazateley KPI monitor. Available: <http://www.kpi-monitor.ru/sorutions/balanced-scorecard/> (Accessed: 23.05.2019). (rus)
10. Karty pokazateley prognoz platform. Available: http://help.prognoz.com/8.0/ru/mergedProjects/UiScoreCard/uiscocard_title.htm/ (Accessed: 23.05.2019). (rus)
11. **Koehler G.** Balanced scorecards in power BI. 2018. January. Available: <https://powerbi.microsoft.com/ru-ru/blog/balanced-scorecards/> (Accessed: 23.05.2019). (rus)
12. **Nah F.** A study on tolerable waiting time: how long are web users willing to wait? *Behaviour and Information Technology*, 2004, Vol. 23, No. 3, Pp. 153–163.
13. **Nielsen J.** Website response times. 2010. June. Available: <https://www.nngroup.com/articles/website-response-times/> (Accessed: 23.05.2019).
14. **Rigby D., Bilodeau B.** Management tools and trends 2015. June. Available: http://www2.bain.com/management_tools/BainTopTenTools/2017/default.asp/ (Accessed: 23.05.2019).
15. **Khedkar S. Atul, Shinde S.K.** Customer review analytics for business intelligence. *2018 IEEE International Conference on Computational Intelligence and Computing Research*, Madurai, India, 2018, Pp. 1–5.
16. **Kamimori S., Ogata S., Kaijiri K.** Automatic method of generating a web prototype employing live interactive widget to validate functional usability requirements. *2015 3rd International Conference on Applied Computing and Information Technology / 2nd International Conference on Computational Science and Intelligence*, Okayama, 2015, Pp. 8–13.
17. **Gavrilova T.A., et al.** On a method of visual models classification. *Business Informatics*, 2013, No. 4 (26), Pp. 21–34. (rus)
18. **Shebotinov A.A., Kolychev V.D.** Business intelligence visualization of strategic KPI organizations using of business intelligence tools. *Skif. Questions of Student Science*, 2018, No. 4 (20). (rus)
19. **Kolychev V.D., Rumyantsev V.P.** Visual models' system of project management. *Scientific Visualization. Electronic Journal*, 2014, Vol. 6, No. 3, Pp. 14–54. (rus)
20. **Tsvetkov V.Ya.** Visual simulation in decision support system. *International Magazine of Applied and Fundamental Researches*, 2016, No. 10-1, Pp. 13–17. (rus)

Received 08.11.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

ПОПОВ Сергей Геннадьевич

POPOV Sergey G.

E-mail: popovserge@spbstu.ru

ПОНОМАРЕВА Елена Всеволодовна

PONOMAREVA Elena V.

E-mail: eponomareva@ibs.ru

САМОЧАДИН Александр Викторович

SAMOCHADIN Alexander V.

E-mail: samochadin@gmail.com

DOI: 10.18721/JCSTCS.12402
УДК 621.396.7

ЧАСТОТНО-НЕЗАВИСИМЫЙ СПОСОБ ЦИФРОВОЙ КАЛИБРОВКИ КАНАЛОВ АЦП УСТРОЙСТВ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ

А.С. Фатеева

Омский государственный технический университет,
Омск, Российская Федерация

Приведено описание метода временного уплотнения сигналов при параллельном включении нескольких АЦП. Способ цифровой калибровки каналов базируется на усовершенствованном методе наименьших квадратов. Для регулировки смещения опорного напряжения и коэффициента усиления используется алгоритм, основанный на рекурсивном методе наименьших квадратов. Преимущества предлагаемого способа на основе итеративного алгоритма – высокое быстродействие и высокая точность настройки каналов АЦП по сравнению со способом, основанным на методе наименьших квадратов. Для реализации процедуры «усреднения» предложен численный метод интегрирования Эйлера. Предлагаемый способ калибровки частотно-независимый, имеющий одинаковую эффективность во всей полосе входного сигнала. Представлены результаты имитационного моделирования, включающие в себя спектральный анализ выходного сигнала до и после калибровки.

Ключевые слова: цифровой способ, калибровка каналов, рекурсивный метод, итеративный алгоритм, численный метод, интегрирование Эйлера.

Ссылка при цитировании: Фатеева А.С. Частотно-независимый способ цифровой калибровки каналов АЦП устройств цифровой обработки сигналов // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 20–28. DOI: 10.18721/JCSTCS.12402

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

FREQUENCY-INDEPENDENT METHOD OF DIGITAL CALIBRATION OF ADC CHANNELS DIGITAL SIGNAL PROCESSING DEVICES

A.S. Fateeva

Omsk State Technical University,
Omsk, Russian Federation

In the work, a description is given of a method for temporarily compressing signals when several ADCs are connected in parallel. The digital channel calibration method is based on an advanced least-squares method. An algorithm based on the recursive least squares method is used to adjust the offset of the reference voltage and gain. An advantage of the proposed method on the basis of an iterative algorithm is high speed and high



accuracy of tuning the ADC channels in comparison with the method based on the least squares method. To implement the «averaging» procedure, a numerical method for Euler integration is proposed. The proposed calibration method is frequency independent and has the same efficiency in the entire input signal band. The results of simulation modeling are presented, including a spectral analysis of the output signal before and after calibration.

Keywords: digital receiving device, receiver bandwidth, timedivision multiplexing, jitter, aperture error.

Citation: Fateeva A.S. Frequency-independent method of digital calibration of ADC channels digital signal processing devices. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 20–28. DOI: 10.18721/JCSTCS.12402

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Введение

В современных устройствах цифровой обработки сигналов широкое распространение получили методы расширения полосы пропускания приёмных трактов, основанные на параллельном включении нескольких аналогово-цифровых преобразователей (АЦП). Параллельное включение нескольких АЦП позволяет использовать метод временного уплотнения сигналов для расширения полосы пропускания приёмного тракта устройств цифровой обработки сигналов [1]. Применение временного уплотнения сигналов позволяет достигать высокого быстродействия приёмных трактов: до 20 Гвыб/с.

Однако при практической реализации у данного метода возникают существенные недостатки, связанные с разбросом параметров реальных АЦП, неконтролируемыми временными задержками трактов, апертурной погрешностью АЦП и т. д. [2]. В конечном итоге это может сказаться на качестве приёма сигналов: на соотношении сигнал/шум и коэффициента искажения (SINAD). Для того чтобы компенсировать большинство из представленных недостатков используют различные способы калибровки каналов АЦП [3]. В последнее время широкое распространение получили цифровые способы калибровки смещения опорного напряжения, коэффициента усиления и временной задержки между каналами [4, 5]. Но существующие способы цифровой калибровки предполагают выполнение процедуры калибровки в специальном режиме, когда приём реального

сигнала не осуществляется, и при этом время, затраченное на выполнение калибровки, снижает степень готовности устройств цифровой обработки сигналов [6, 7]. В некоторых случаях степень готовности устройства цифровой обработки сигналов – это ключевой показатель всей системы связи. Поэтому ускорение процедуры калибровки каналов АЦП является одной из первоочередных задач при создании устройств цифровой обработки сигналов, использующих метод временного уплотнения.

Временное уплотнение сигналов при параллельном включении нескольких АЦП

Структурная схема метода, реализующего временное уплотнение сигналов при параллельном включении четырёх АЦП, представлена на рис. 1.

В данной структурной схеме входной аналоговый сигнал одновременно поступает на вход четырёх АЦП (ADC1...4). Каждый из АЦП подключен к синтезатору тактовых сигналов таким образом, чтобы обеспечить поочередное смещение начальных фаз тактовых сигналов на 90°. Это позволяет обеспечить синхронность временного уплотнения сигналов в цифровом виде и восстановить цифровой сигнал с помощью цифрового мультимплексора, частота дискретизации которого в четыре раза больше частоты преобразования АЦП. Таким образом достигается четырёхкратное расширение полосы пропускания для устройства цифровой обработки сигналов. Очевидно, что качество принимаемого

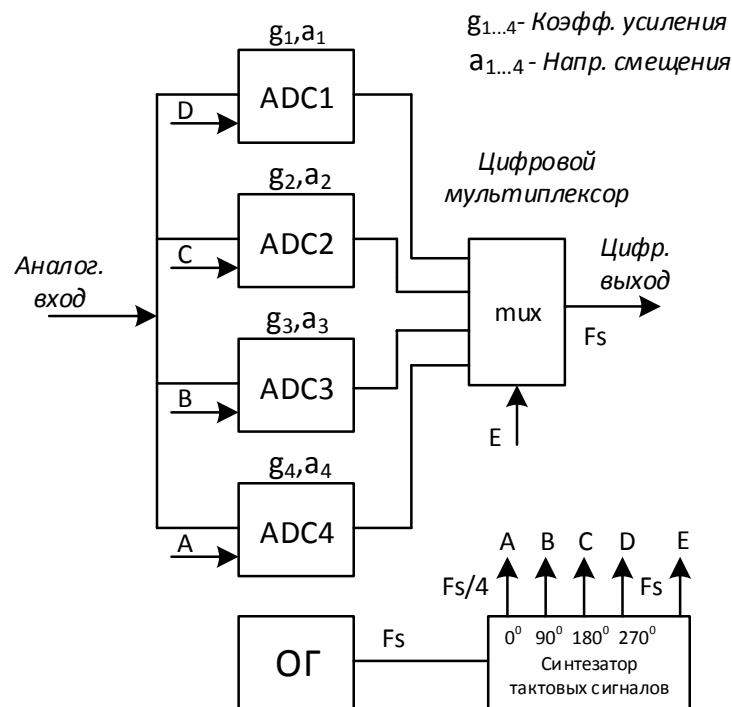


Рис. 1. Структурная схема метода временного уплотнения сигналов при параллельном включении четырёх АЦП

Fig. 1. Block diagram of the method of temporary compaction of signals at parallel inclusion of 4 ADC

сигнала будет зависеть от идентичности параллельных каналов АЦП. Различие в параметрах каналов АЦП будет выражаться в ухудшении соотношения сигнал/шум и коэффициента искажения (SINAD).

Метод наименьших квадратов для корректировки смещения опорного напряжения и коэффициента усиления

Смещение опорного напряжения АЦП нередко связано с разбросом параметров встроенных операционных усилителей и компараторов. Кроме того, смещение опорного напряжения АЦП может возникнуть из-за асимметрии каскадов преобразования АЦП и связано с погрешностями при производстве микросхем на фабрике.

Различие в коэффициентах усиления нескольких АЦП обусловлено паразитными ёмкостями и индуктивностями операционных усилителей, а также отличием параметров входных аналоговых конденсаторов.

Чтобы найти смещение опорного напряжения необходимо выбрать опорное напряжение одного из АЦП в качестве эталона. Тогда усреднённое среднеквадратичное отклонение цифрового сигнала с выхода параллельных АЦП относительно эталонного АЦП будет определять смещение опорного напряжения [8]. Различие в коэффициентах усиления параллельных АЦП можно найти аналогичным способом, за исключением того, что вместо отклонения цифрового сигнала потребуется значение модуля этого сигнала на выходе каждого из АЦП.

Применение итеративного алгоритма на основе метода наименьших квадратов (LMS) позволяет находить и компенсировать смещение опорного напряжения и различие коэффициентов усиления параллельных АЦП. Для этого необходимо выделить один из параллельных АЦП в качестве опорного (например, первый АЦП). Тогда итеративный алгоритм на основе ме-



тогда наименьших квадратов для нахождения напряжения смещения параллельно включенных АЦП будет выглядеть следующим образом:

$$a_i(n) = a_i(n-1) + \mu(\hat{x}(n-1) - x_i(n-1)), \quad (1)$$

где $x_i(n)$ – цифровой выход с опорного (первого) АЦП; $\hat{x}(n)$ – цифровой выход с одного из параллельных АЦП; μ – шаг адаптации; a_i – смещение опорного напряжения одного из параллельных АЦП относительно опорного АЦП.

В свою очередь алгоритм для нахождения коэффициентов усиления параллельно включенных АЦП можно представить следующим выражением:

$$k_i(n) = k_i(n-1) + \mu_K(\hat{x}(n-1) - x_i(n-1)), \quad (2)$$

где $x_i(n)$ – цифровой выход с опорного (первого) АЦП; $\hat{x}(n)$ – цифровой выход с одного из параллельных АЦП; μ_K – шаг адаптации; k_i – параметр, определяющий различие коэффициентов усиления одного из параллельных АЦП относительно опорного АЦП. При этом параметр k_i определяется следующим соотношением:

$$k_i = \frac{g_1}{g_i}, \quad (3)$$

где g_1 – коэффициент усиления опорного АЦП (первого); g_i – коэффициент усиления одного из параллельных АЦП.

Полученные с помощью выражений (1) и (2) значения a_i и k_i необходимо «усреднять» для повышения точности и скорости калибровки. Для усреднения значений a_i и k_i , как правило, используют традиционные методы, основанные на накоплении большого количества данных, и методы, основанные на экспоненциальном усреднении [9]. Использование традиционных методов усреднения требует большого количества вычислительных ресурсов. Экспоненциальное усреднение не требует большого количества вычислительных ресурсов, однако предполагает использование двух параметров для усреднения (прямого и обратного коэффициентов усиления). При этом для нахождения обратного коэффи-

циента усиления требуется использование операции деления.

Численный метод интегрирования Эйлера для усреднения данных

Альтернативным способом «усреднения» коэффициентов, полученных с помощью выражений (1) и (2), может быть способ на основе численного метода интегрирования Эйлера. Основой интегрирования является классическая формула Эйлера:

$$y(t + \Delta t) = y(t) + \Delta t \cdot f(x(t), y(t), t) + \varepsilon, \quad (4)$$

где $y(t)$ – текущее состояние; Δt – шаг приращения; $f(x(t), y(t), t)$ – скорость приращения.

На основе формулы (3) можно получить выражение для интегрирования в дискретной форме. В зависимости от того, какие входные данные будут учитываться при вычислении результата, метод Эйлера может быть прямым или обратным [10]. Для реализации функции «усреднения» удобнее использовать прямой метод.

Прямой метод Эйлера описывается следующим выражением:

$$y(n) = y(n-1) + P[t(n) - t(n-1)]x(n-1), \quad (5)$$

где $y(n)$ – текущее значение функции; P – коэффициент усиления (определяет усиление для шага приращения); $x(n)$ – входные данные.

Тогда для «усреднения» смещения опорного напряжения одного из параллельных АЦП относительно опорного АЦП (a_i) можно записать следующее выражение:

$$A_i(n+1) = (A_i(n) + P[t(n) - t(n-1)]A_i(n-1)), \quad (6)$$

где $A_i(n) = (a_i(n) - a_i(n-1))$.

На практике реализация алгоритма «усреднения» на основе выражения (5) потребует трёх сумматоров и двух умножителей. Отличительной особенностью будет являться то, что для изменения шага приращения (скорости усреднения) используется один коэффициент P , что в значительной степени упрощает настройку параметров системы калибровки АЦП.

Аналогичным образом можно описать функцию «усреднения» коэффициентов усиления параллельно включенных АЦП:

$$K_i(n+1) = (K_i(n) + P[t(n) - t(n-1)]K_i(n-1)), \quad (7)$$

где $K_i(n) = (k_i(n) - k_i(n-1))$.

Стоит отметить, что использование больших значений для коэффициента P приводит к увеличению ошибки ε и, соответственно, точности калибровки АЦП в целом.

Имитационное моделирование предлагаемого способа корректировки смещения опорного напряжения и коэффициента усиления

Для исследования и анализа эффективности предлагаемого способа для корректировки смещения опорного напряжения и коэффициента усиления разработана специализированная имитационная модель. Данная модель включает в себя четыре па-

раллельно включенных 12-битных АЦП, синтезатор тактовых частот, цифровой мультиплексор, опорный генератор и блок корректировки коэффициента усиления и напряжения смещения. Структурная схема модели представлена на рис. 2.

В данной структурной схеме блок корректировки коэффициента усиления и напряжения смещения реализует алгоритмы цифровой обработки сигналов в соответствии с формулами (1)–(3) и (5), (6). В качестве опорного АЦП был выбран первый АЦП. Исходные данные для моделирования представлены в табл. 1.

На рис. 3 изображён спектр сигнала на выходе первого АЦП при воздействии синусоидального сигнала частотой 10 МГц.

Спектральные характеристики сигнала на выходе первого АЦП:

- соотношение сигнал/шум (SNR) 62,0 дБ;

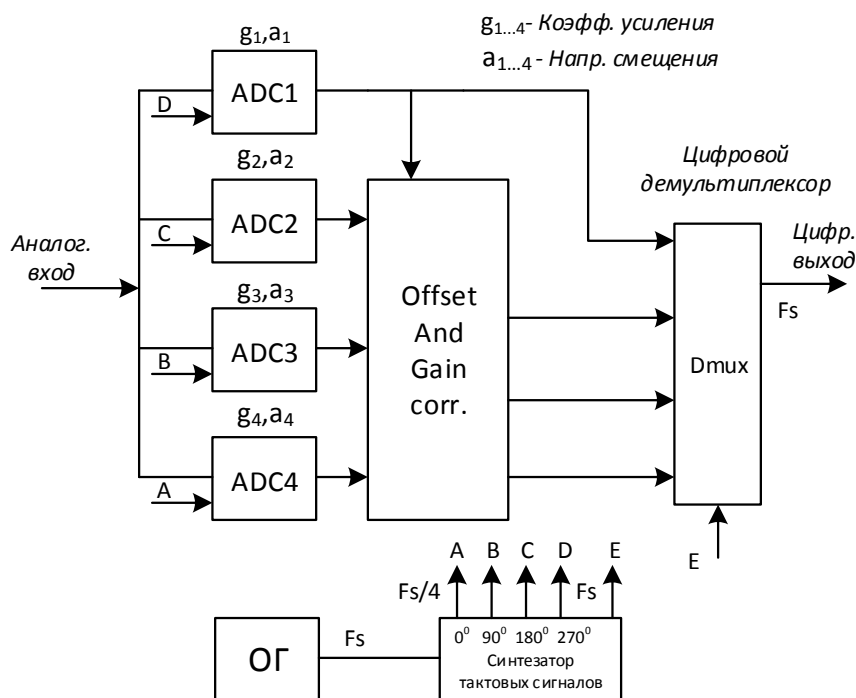


Рис. 2. Структурная схема имитационной модели метода временного уплотнения при воздействии сигнала с блоком коррекции опорного напряжения и коэффициента усиления

Fig. 2. Block diagram of the simulation model of the method of temporary compaction under the influence of a signal with a block of correction of the reference voltage and gain

Таблица 1

Исходные данные для моделирования

Table 1

Initial data for modeling

Номер АЦП	1	2	3	4
Различие опорного напряжения	0	0,005	−0,005	0,001
Различие коэффициентов усиления	1	1,005	0,995	1,01
Частота дискретизации, МГц	125	125	125	125
Разрядность, бит	12	12	12	12

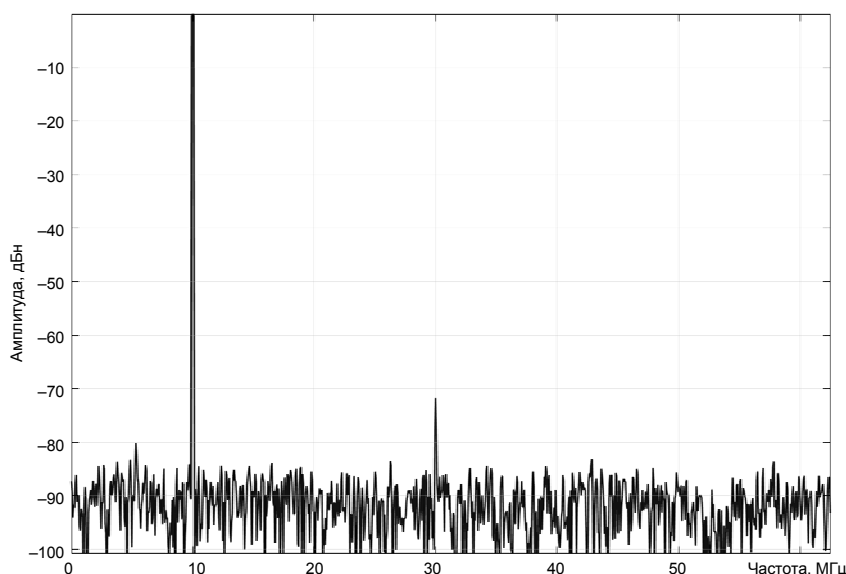


Рис. 3. Спектр сигнала на выходе первого АЦП при воздействии синусоидального сигнала частотой 10 МГц

Fig. 3. The spectrum of the signal at the output of the first ADC when exposed to a sinusoidal signal frequency of 10 MHz

- динамический диапазон, свободный от искажений (SFDR) 71,7 дБ;

- коэффициент гармонических искажений (THD) −71,2 дБ.

Далее выполнялся анализ сигнала на выходе цифрового мультиплексора без включения блока коррекции коэффициента усиления и напряжения смещения. На рис. 4 представлен спектр сигнала на выходе цифрового мультиплексора без калибровки. Как видно из графика, в спектре сигнала присутствуют паразитные компоненты.

Спектральные характеристики сигнала на выходе цифрового мультиплексора (без калибровки):

- соотношение сигнал/шум (SNR) 53,7 дБ;

- динамический диапазон, свободный от искажений (SFDR) 56,3 дБ;

- коэффициент гармонических искажений (THD) −68,3 дБ.

Затем анализировался сигнал на выходе цифрового мультиплексора с предварительно выполненной калибровкой коэффициента усиления и напряжения смещения всех АЦП. Параметры блока коррекции были выбраны следующими:

- шаг адаптации для коррекции смещения опорного напряжения $\mu = 0,1$;

- шаг адаптации для коррекции коэффициента усиления $\mu_K = 0,1$;

- коэффициент для «усреднения» смещения опорного напряжения $PA = 40\,000$;

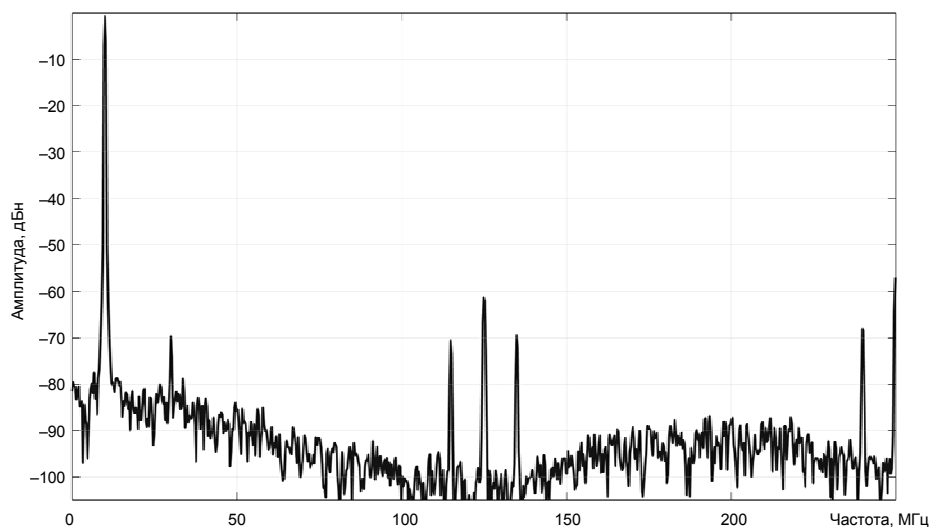


Рис. 4. Спектр сигнала на выходе цифрового мультиплексора при воздействии синусоидального сигнала частотой 10 МГц (без калибровки)

Fig. 4. The spectrum of the signal at the output of the digital multiplexer when exposed to a 10 MHz sinusoidal signal (without calibration)

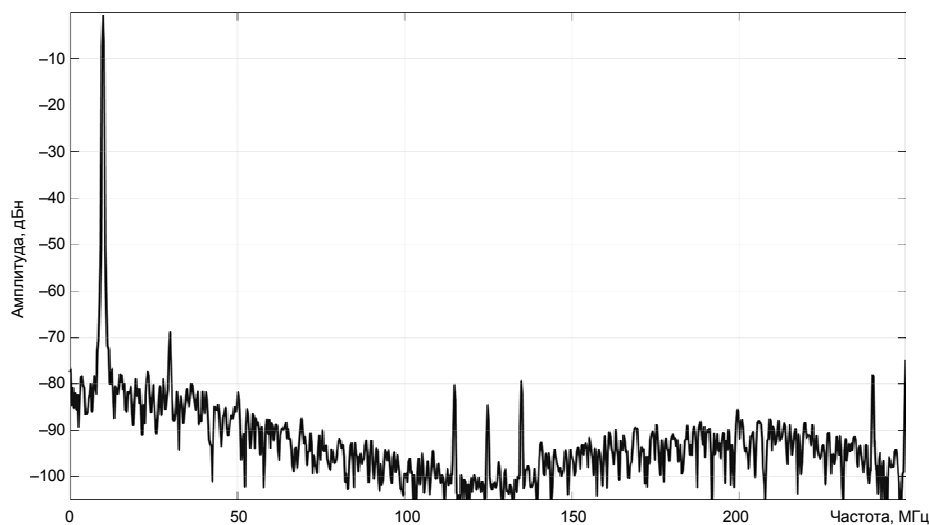


Рис. 5. Спектр сигнала на выходе цифрового мультиплексора при воздействии синусоидального сигнала частотой 10 МГц (с калибровкой)

Fig. 5. The spectrum of the signal at the output of the digital multiplexer when exposed to a sinusoidal signal frequency of 10 MHz (with calibration)

• коэффициент для «усреднения» коэффициента усиления $RK = 40\,000$.

На рис. 5 представлен спектр сигнала на выходе цифрового мультиплексора с откалиброванными АЦП. Как видно из графика, уровень паразитных компонент в

спектре выходного сигнала значительно уменьшился.

Спектральные характеристики сигнала на выходе цифрового мультиплексора с включенным блоком коррекции:

- соотношение сигнал/шум (SNR) 61,3 дБ;



Таблица 2
Спектральные характеристики сигнала при воздействии различных частот

Table 2

Spectral characteristics of the signal when exposed to different frequencies

Частота сигнала, МГц	Спектральные характеристики					
	SNR, дБ		SFDR, дБ		THD, дБ	
	без коррекции	с коррекцией	без коррекции	с коррекцией	без коррекции	с коррекцией
10	53,7	61,3	56,3	68,9	−68,3	−68,2
50	46,5	59,2	51,8	66,7	−51,6	−51,4
150	37,4	55,7	50,6	66,3	−48,7	−48,9
200	34,7	53,6	50,2	66,2	−47,8	−47,7

- динамический диапазон, свободный от искажений (SFDR) 68,9 дБ;
- коэффициент гармонических искажений (THD) −68,2 дБ.

Результаты имитационного моделирования

В табл. 2 представлены спектральные характеристики сигнала (SNR, SFDR, THD) на выходе имитационной модели с откалиброванными и неоткалиброванными АЦП. Параметры блока коррекции: $\mu = 0,1$; $\mu_K = 0,1$; $P_A = 40\,000$; $P_K = 40\,000$.

Исходя из представленных данных, можно сделать вывод о том, что использование предлагаемого способа калибровки для коррекции смещения опорного напряжения и коррекции коэффициента усиления позволяет существенным образом улучшить спектральные характеристики сигнала при использовании метода временного уплотнения сигнала.

Заключение

Использование предлагаемого в данной статье способа калибровки АЦП для метода временного уплотнения сигналов позволяет в значительной мере повысить динамические характеристики сигнала. Применение численного метода интегрирования Эйлера для реализации процедуры «усреднения» позволяет существенно сократить вычислительные ресурсы. В ходе анализа результатов имитационного моделирования было установлено, что предлагаемый способ калибровки является частотно-независимым и одинаково эффективен во всей полосе входного сигнала. Ограничением для использования способа может являться выбор параметров для блока коррекции, т. к. скорость калибровки и точность калибровки — это взаимоисключающие понятия. Чем меньше шаг адаптации μ и больше коэффициент усиления P , тем точнее калибруются АЦП, но при этом существенно возрастает время калибровки.

СПИСОК ЛИТЕРАТУРЫ

1. Black W.C., Hodges D.A. Time-interleaved converter arrays // IEEE J. Solid-State Circuits. 1980. Vol. 15. No. 12. Pp. 1022–1029. DOI: 10.1109/JSSC.1980.1051512
2. Fateeva A., Nikonova G. Digital receiving devices with time-division multiplexing for bandwidth extension // In Proc. of Ural Symp. on Biomedical Engineering, Radioelectronics and Information Technology. Ekaterinburg, Apr. 2019. DOI: 10.1109/USBREIT.2019.8736656
3. Reyes B.T., Tealdi L., Paulina G., et al. A 6-bit 2GS/s CMOS time-interleaved ADC for analysis of mixed-signal calibration techniques // In Proc. of the IEEE 5th Latin American Symp. on Circuits and Systems. Santiago, Chile, Febr. 2014. Pp. 1–4. DOI: 10.1109/TCSI.2016.2636209
4. Jin H., Lee E.K.F. A digital-background calibration technique for minimizing timing-error effects in time-interleaved ADCs // IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.

2000. Vol. 47. No. 7. Pp. 603–613. DOI: 10.1109/82.850419

5. Yin Y., Yang G., Chen H. A novel gain error background calibration algorithm for time-interleaved ADCs // In Proc. of the Internat. Conf. on Anti-counterfeiting, Security, and Identification. Macao, China, 2014. Pp. 42–45. DOI: 10.1049/iet-cds.2018.5194

6. Wei H.G., Zhang P., Sahoo B.D., Razavi B. An8 Bit4GS/s 120mWCMOS ADC // IEEE Journal of Solid-State Circuits, 2014, Vol. 49, No. 8, Pp. 1751–1761. DOI: 10.1109/JSSC.2014.2313571

7. Chen H., Yin Y. A low complexity all-digital background calibration technique for time-interleaved ADCs. Hindawi Publishing Corpora-

tion. VLSI Design, 2016. Article ID 6475932 // URL: <http://dx.doi.org/10.1155/2016/6475932>

8. Eklund J., Gustafsson F. Digital offset compensation of time interleaved ADC using random chopper sampling // In IEEE Symp. Circuits Syst. May 2000. Pp. 447–450 // URL: https://doi.org/10.1007/978-90-481-9725-5_5

9. Yin Y., Li J., Chen H. A digital background calibration algorithm of time-interleaved ADC // In Proc. of the Internat. Conf. on Anti-counterfeiting, Security, and Identification. Macao, China, 2014, Pp. 1–4. DOI: 10.1109/TCSI.2008.2001830

10. Butcher J.C. Numerical methods for ordinary differential equations. New York: John Wiley & Sons, 2010. DOI: 10.1007/s11075-009-9285-0

Статья поступила в редакцию 23.10.2019.

REFERENCES

1. Black W.C., Hodges D.A. Time-interleaved converter arrays. *IEEE Journal Solid-State Circuits*, 1980, Vol. 15, No. 12, Pp. 1022–1029. DOI: 10.1109/JSSC.1980.1051512

2. Fateeva A., Nikonova G. Digital receiving devices with time-division multiplexing for bandwidth extension. In *Proceedings of Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology*, Ekaterinburg, Russia, Apr. 2019. DOI: 10.1109/USBREIT.2019.8736656

3. Reyes B.T., Tealdi L., Paulina G., et al. A 6-bit 2GS/s CMOS time-interleaved ADC for analysis of mixed-signal calibration techniques. In *Proceedings of the IEEE 5th Latin American Symposium on Circuits and Systems*, Santiago, Chile, Febr. 2014, Pp. 1–4. DOI: 10.1109/TCSI.2016.2636209

4. Jin H., Lee E.K.F. A digital-background calibration technique for minimizing timing-error effects in time-interleaved ADCs. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, Jul. 2000, Vol. 47, No. 7, Pp. 603–613. DOI: 10.1109/82.850419

5. Yin Y., Yang G., Chen H. A novel gain error background calibration algorithm for time-interleaved ADCs. In *Proceedings of the International Conference on Anti-counterfeiting, Security, and Identification*,

Macao, China, December 2014, Pp. 42–45. DOI: 10.1049/iet-cds.2018.5194

6. Wei H.G., Zhang P., Sahoo B.D., Razavi B. An8 Bit4GS/s 120mWCMOS ADC. *IEEE Journal of Solid-State Circuits*, 2014, Vol. 49, No. 8, Pp. 1751–1761. DOI: 10.1109/JSSC.2014.2313571

7. Chen H., Yin Y. A low complexity all-digital background calibration technique for time-interleaved ADCs. Hindawi Publishing Corporation, Vol. 2016, VLSI Design, 2016, Article ID 6475932. Available: <http://dx.doi.org/10.1155/2016/6475932>

8. Eklund J., Gustafsson F. Digital offset compensation of time interleaved ADC using random chopper sampling. In IEEE Symp. Circuits Syst., May 2000, Pp. 447–450. Available: https://doi.org/10.1007/978-90-481-9725-5_5

9. Yin Y., Li J., Chen H. A digital background calibration algorithm of time-interleaved ADC. In *Proceedings of the International Conference on Anti-counterfeiting, Security, and Identification*, Macao, China, Dec. 2014, Pp. 1–4. DOI: 10.1109/TCSI.2008.2001830

10. Butcher J.C. *Numerical methods for ordinary differential equations*. New York: John Wiley & Sons, 2010. DOI: 10.1007/s11075-009-9285-0

Received 23.10.2019.

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

ФАТЕЕВА Анна Сергеевна

FATEEVA Anna S.

E-mail: fateewa16@rambler.ru

DOI: 10.18721/JCSTCS.12403
УДК 004.4'416:004.382.2

МЕТОД ГИБРИДНОГО НЕОДНОРОДНОГО ТАЙЛИНГА ДЛЯ АРХИТЕКТУР СУПЕРКОМПЬЮТЕРОВ С МНОГОУРОВНЕВОЙ ИЕРАРХИЕЙ ПАМЯТИ

А.В. Левченко

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Предложен метод разбиения операций вычислительного алгоритма, основанный на выполнении преобразований в рамках полиэдральной модели компиляции в сочетании с классическим иерархическим параметризованным тайлингом для последующего преобразования кода, представленного в виде абстрактного синтаксического дерева. Разработана последовательность преобразований, позволяющих осуществить приоритизацию локальности на полиэдральной стадии, а на стадии многоуровневого тайлинга — сгенерировать тайлы на основе моделей для отображения на иерархическую архитектуру памяти. Построены производные алгоритмы акогерентного неоднородного тайлинга, расширяющие возможности предложенного метода посредством подстановки вариантов алгоритмов выбора формы и размера тайлов. Получены экспериментальные результаты, позволяющие оценить преимущества предложенного метода в контексте достижения переносимости производительности вычислительных алгоритмов на многомашинные макроузлы с глобально адресуемой когерентной памятью с неоднородным доступом.

Ключевые слова: иерархический тайлинг, компиляция, локальность, многоуровневая иерархия памяти, оптимизации цикла, переносимость производительности, полиэдральная модель.

Ссылка при цитировании: Левченко А.В. Метод гибридного неоднородного тайлинга для архитектур суперкомпьютеров с многоуровневой иерархией памяти // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 29–44. DOI: 10.18721/JCSTCS.12403

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

A HYBRID NON-UNIFORM LOOP TILING FOR SUPERCOMPUTERS WITH DEEP MEMORY HIERARCHIES

A.V. Levchenko

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

Non-uniform nature of multi-level memory architectures of modern supercomputers represents a notably underestimated issue in design of loop transformation algorithms. The imperfect compiler support for architectural features of

deep memory hierarchies results in insufficient data locality, which in turn is an obstacle to achieving performance portability for a wide range of important computational kernels like iterated stencils or sparse-matrices. The major contribution of this paper is an algorithmic skeleton for hybrid non-uniform loop tiling. The proposed approach combines locality-enhancing features of polyhedral compilation framework with capability of non-uniformity effects modeling via hierarchical parameterized tiling strategy performed in a canonical syntactic manner. The polyhedral stage focuses on spatial and temporal locality prioritization along with end-to-end optimization pipeline. At the syntactic stage the parameterized loop tiling strategy allows an automatic definition of tiled loop characteristics to map it according to the hierarchical memory architecture. The tiles with various parameters like size, shape and form can be generated through the novel permutational target-specific algorithms. As a result, the variants of acoherent non-uniform loop tiling algorithm were designed on the basis of the proposed approach to evaluate the permutational techniques of tile size and tile shape selection. The early-stage experimental results are presented to show the effects of hybrid non-uniform tiling on data locality to deliver near-optimal performance portability for representative benchmarks across deepening memory hierarchies of multi-machine macronodes.

Keywords: compiler optimizations, data locality, hierarchical tiling, loop optimizations, multi-level memory hierarchy, performance portability, polyhedral model.

Citation: Levchenko A.V. A hybrid non-uniform loop tiling for supercomputers with deep memory hierarchies. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 29–44. DOI: 10.18721/JCSTCS.12403

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Перспективным методом улучшения свойств вычислительных алгоритмов для эффективного использования многоуровневой иерархии памяти суперкомпьютеров является применение математической модели преобразований циклов, известной как полиэдральная модель компиляции. К замечательным свойствам полиэдрального подхода относится возможность реализации новых алгоритмов компиляции посредством композиции множества сложных автоматических преобразований-примитивов модели в виде единой алгебраической операции с возможностью последующего моделирования преобразований при помощи методов оптимизации [1].

В частности, одним из примитивов полиэдральной модели компиляции является метод разделения операций программной реализации вычислительного алгоритма на подмножества для реорганизации порядка вычислений: тайлинг. Тайлинг — фундаментальное преобразование для улучшения

повторного использования данных при осуществлении доступа к регионам памяти, к которым доступ осуществлялся недавно, либо к соседствующим с ними регионам, что обуславливает временную и пространственную локальность вычислительных алгоритмов в неоднородной иерархической структуре памяти современных суперкомпьютеров [2].

Альтернативой тайлингу, выполняемому в рамках полиэдральной модели компиляции, является иерархический параметризованный тайлинг, основанный на применении корректных моделей параметров многоуровневой архитектуры памяти и использующий классический вариант преобразований вложенных циклов кода, представленного в виде абстрактного синтаксического дерева.

Несмотря на ряд полученных недавно результатов в изучении свойств полиэдрального и синтаксического подходов [3–8], на практике оказывается, что, будучи отдельно взятыми, эффекты этих подходов явля-



ются недостаточными [9]. Причина состоит в том, что существующие методы компиляции в должной степени не учитывают особые уровни неоднородности архитектуры и локальности доступа к памяти вычислительных алгоритмов при выполнении на реальных высокопроизводительных системах с глубокой иерархической памятью. В связи с этим для данных целевых архитектур должна быть решена задача построения усовершенствованных комбинированных методов тайлинга посредством использования преимуществ полиэдрального и синтаксического подходов.

Цель настоящего исследования — разработка метода гибридного неоднородного тайлинга, на основе которого будут построены алгоритмы преобразования циклов для достижения переносимости производительности вычислений на архитектуры суперкомпьютеров с многоуровневой иерархией памяти.

Обзор литературы

Сформулируем особые требования к методам преобразований циклов, на основе которых будет выполнен анализ предметной области. Эти требования основаны на ряде недавних оценок положения дел в области высокопроизводительных вычислений [10–12] и, в частности, на прогнозе [13], согласно которому одним из принципов построения некоторых типов суперкомпьютеров экза- и зеттафлопсного уровня производительности станет наличие логически неделимой, глобально адресуемой многоуровневой архитектуры памяти, полученной путем объединения памяти вычислительных узлов, что неизбежно обусловит неудовлетворительную локальность широкого спектра научных вычислительных алгоритмов. Поэтому локальность становится главным фактором в модели для переносимости производительности [14] на системы данного класса. При этом системы с глобально адресуемой когерентной памятью с неоднородным доступом могут рассматриваться в качестве прототипа для гипотетических машин экзафлопсного

уровня. Это следует из результатов работ [15–19], общей чертой которых является рассмотрение в таком качестве далеко не столь крупномасштабных систем с неоднородным доступом к памяти.

Таким образом, учитывая особенности целевой архитектуры, целесообразно сформулировать следующие требования к уже существующим методам автоматического преобразования циклов. Первое требование состоит в необходимости моделирования эффектов пространственной и временной локальности посредством применения алгоритмов планирования в полиэдральной модели компиляции [4]. Второе требование включает моделирование неоднородности посредством иерархического тайлинга. Этот метод предполагает параметризованное определение свойств атомарно выполняемых операций для отображения на архитектуру памяти в соответствии с конфигурацией кэш-памяти, узлов NUMA, буфера ассоциативной трансляции. Наконец, третье требование касается необходимости экспериментальной оценки производительности результирующего преобразованного кода с использованием реальной системы с многоуровневой иерархией памяти.

Полиэдральная модель компиляции рассматривается в ряде недавних работ [20–24] как алгоритмический аппарат, в рамках которого тайлинг может быть применен для улучшения локальности. Важным примером представляется работа [4], в которой предложен перспективный метод оптимизации пространственной и временной локальности систем с глубокой иерархией памяти посредством решения задачи целочисленного линейного программирования с использованием лексикографически минимального вектора пространства поиска. Результаты расширенного исследовательского отчета [25], предшествующего указанной работе, свидетельствуют о превосходстве предложенного алгоритма над существующими полиэдральными, синтаксическими и гибридными аналогами, воплощенными в современных оптимизато-

рах. Однако существуют и некоторые общие для полиэдральных техник ограничения, касающиеся преобразования статических областей управления, подробно описанные в [26], так же как и ограничения, связанные с возможным наличием скалярных зависимостей, рассмотренные в [27], и ряд других вопросов [28]. Кроме того, в [29] отмечена неудовлетворительная степень масштабируемости полиэдральной модели. На практике следствием перечисленных особенностей является недостаточность суперкомпьютерных полиэдральных тестов производительности: экспериментальные исследования обычно ограничиваются использованием набора программ PolyBench/C, примером чего являются многие из перечисленных работ.

Применение исключительно полиэдрального подхода представляется недостаточным для высокопроизводительных систем с глубокой иерархической памятью. В связи с этим в работе [30] была успешно проверена концепция комбинированного применения преобразований полиэдрального представления для локальности с последующими синтаксическими преобразованиями для повторного использования данных, межтайловой параллелизации и внутритайловой векторизации. Некоторые современные варианты синтаксического тайлинга рассмотрены в обзоре [2]. Среди перспективных возможностей можно допустить подстановку в разрабатываемый метод алгоритмов выбора размера тайла, предложенных в [31, 32]. В зависимости от специфических особенностей целевой архитектуры могут также применяться аналогичные процедуры для трафаретных вычислений, описанные в [33–35]. Наконец, решению проблемы неоднородности адресовано исследование [36], в котором предложен метод синтаксического неоднородного тайлинга, основанный на моделировании абстрактных уровней иерархии памяти. При этом данный метод не фокусируется на аспектах локальности доступа.

Обзор литературы привел к выводу о том, что рассмотренные подходы в полной

мере не удовлетворяют всем сформулированным требованиям. В настоящем исследовании при разработке метода гибридного неоднородного тайлинга предлагается выполнить все требования за счет использования преимуществ существующих подходов.

Метод гибридного неоднородного тайлинга

Предварительные сведения. Метод гибридного неоднородного тайлинга основан на выполнении преобразований в рамках полиэдральной модели компиляции в сочетании с синтаксическим иерархическим тайлингом. Главная цель разработки метода состоит в достижении переносимости производительности посредством улучшения локальности, важного свойства научных вычислительных алгоритмов. На основе концепции [30] разработана последовательность необходимых преобразований, позволяющих осуществить приоритизацию пространственно-временной локальности на полиэдральной стадии. Далее многоуровневый тайлинг осуществляется для последующего преобразования кода, представленного в виде абстрактного синтаксического дерева с генерацией тайлов необходимой размерности. Тайлы генерируются на основе уточненных моделей для отображения на неоднородную иерархическую архитектуру памяти.

Метод предназначен для улучшения локальности научных вычислительных алгоритмов, в частности, трафаретных вычислений (stencil computations). К данному классу относятся многосеточные методы и их компоненты: симметричный метод Гаусса-Зейделя, метод Гаусса-Якоби, метод последовательной верхней релаксации (SOR) и ряд других методов. Производительность трафаретных вычислений имеет значение для научных вычислений, поскольку данный класс задач лежит в основе решателей уравнений с частными производными. Так, репрезентативный тест производительности суперкомпьютеров HPCG (High Performance Conjugate Gradients) основан на решении системы линейных алгебраических уравнений с



разреженной симметричной матрицей методом сопряженных градиентов с предобуславливателем Гаусса-Зейделя.

Важной особенностью трафаретных вычислений является ограничение пропускной способности памяти ввиду низкого соотношения числа операций с плавающей точкой на байт информации, считываемой из памяти. Поэтому повышение производительности приложений этого класса неизбежно требует снижения перемещения данных из неоднородной памяти посредством увеличения их повторного использования. Таким образом, улучшение локальности трафаретных вычислений будет означать улучшение производительности для широкого спектра вычислительных алгоритмов при переносе на системы с глубокой иерархической памятью.

Описание метода. Метод гибридного неоднородного тайлинга предусматривает выполнение необходимых преобразований на протяжении двух стадий. На первой стадии решается задача улучшения локальности посредством полиэдральных преобразований. Вообще говоря, для иерархии памяти вычислительного узла цепочка преобразований могла быть ограничена только первой стадией, и тайлинг был бы осуществлен в рамках полиэдральной модели. Несмотря на это, именно вторая стадия является ядром предлагаемого метода.

Основная задача здесь состоит в обеспечении более эффективного использования глобально адресуемой когерентной памяти с неоднородным доступом. Для этого реализуется тайлинг, основанный на моделировании неоднородного доступа на разных уровнях иерархии памяти. В полной версии метода на первой стадии тайлинг только подготавливается посредством перестановки циклов и, в перспективе, посредством слияния циклов, но осуществляется уже на второй стадии. При этом применяется тайлинг в рамках синтаксического подхода как более простой и эффективный способ моделирования ряда аспектов архитектуры памяти, что неоднократно было отмечено в работах

[30, 36]. Следует отметить, что в этом отношении потенциал метода раскрыт далеко не в полной мере. Рассмотрим возможности применения тайлинга в рамках предлагаемого метода.

Первая стадия метода предполагает наличие возможности применения тайлинга в рамках математического подхода полиэдральной компиляции. Тайлинг применяется к статическим областям управления (Static Control Parts — SCoPs), которые представляют собой максимальное множество последовательных выражений, где границы цикла и условные операторы являются аффинными функциями окружающих итераторов циклов и параметров, константные значения которых неизвестны на этапе компиляции [1].

После того, как статические области управления будут обнаружены в промежуточном представлении программы, происходит их преобразование в полиэдральное представление. Далее тайлинг осуществляется наряду с прочими преобразованиями за один проход компилятора как единая алгебраическая операция. Это дает возможность приоритизировать улучшение пространственной и временной локальности, моделируя тайлинг посредством численных методов оптимизации.

Как один из примитивов полиэдральной модели компиляции тайлинг применяется в полиэдральном представлении, основанном на геометрическом представлении итераций цикла, описанных в пространстве итераций в виде целочисленного полиэдроид. Полиэдроид охватывает множество точек в векторном пространстве $\mathbb{Z}^n$, удовлетворяющих неравенству

$$\mathcal{D} = \{x \mid x \in \mathbb{Z}^n, Ax + a \geq 0\}, \quad (1)$$

где x — вектор значений счетчика цикла; A — константная матрица; a — константный параметрический вектор.

Область итераций выражения является множеством его итерационных векторов, состоящих из значений итераторов цикла, содержащих выражение, и подмножеством

полного возможного пространства итераций $\mathcal{D} \subseteq \mathbb{Z}^n$ [1]. Область итераций $\mathcal{D}$ выражения S представлена в качестве полиэдроид:

$$\mathcal{D}_S(\vec{p}) = \left\{ \vec{i}_S \mid D_S \begin{pmatrix} \vec{i}_S \\ \vec{p} \\ 1 \end{pmatrix} \geq \vec{0} \right\}, \quad (2)$$

где $\vec{p}$ — вектор параметров; $\vec{i}_S \in \mathbb{Z}^{\dim(\vec{i}_S)}$ — вектор итераций выражения S ; $D_S \in \mathbb{Z}^{m_{D_S} \times (\dim(\vec{i}_S) + \dim(\vec{p}) + 1)}$ — целочисленная матрица с константами, где m_{D_S} — число условий.

Данный подход позволяет применять методы линейной алгебры и линейного программирования для анализа и преобразования итерационных пространств в полиэдральном представлении. Таким образом, программа представлена в виде набора полиэдров, соответствующих множеству выражений $S_i \in S$, заключенных в одном или нескольких циклах. Порядок экземпляров выражений меняется посредством планирования и определяется отношениями планирования, когда каждому выражению присваивается логическая дата $\vec{t}_S$:

$$\Theta_S(\vec{p}) = \left\{ \vec{i}_S \rightarrow \vec{t}_S \mid T_S \begin{pmatrix} \vec{i}_S \\ \vec{t}_S \\ \vec{p} \\ 1 \end{pmatrix} \geq \vec{0} \right\}, \quad (3)$$

где T_S — целевое выражение [37].

Тайлинг осуществляется при условии, что функции планирования не имеют зависимостей, и порядок этих функций можно менять без искажения первоначальной семантики программы. Таким образом, при обнаружении групп взаимозаменяемых (допускающих перестановку) функций планирования осуществляется полное планирование и выполняется тайлинг. Преобразование посредством вставки копии взаимозаменяемой группы соответствующей размерности осуществляется непосредственно перед группой с установленной новой размерности. При этом суще-

ствует возможность варьирования формы и размеров тайла. Для приоритизации локальности применяется метод лексикографической минимизации компонентов вектора пространства поиска [25], где для построения целевой функции используются отношения близости (пространственные и временные), т. к. расстояние между связанными экземплярами минимизируется для улучшения локальности обращения к памяти. Поскольку ожидается, что улучшение временной локальности даст больший прирост производительности, чем улучшение пространственной локальности, то в первую очередь минимизируются временные отношения близости, которые размещаются первыми.

Преобразования перестановки циклов и, в перспективе, слияние циклов используются для подготовки тайлинга на основе аналитической модели доступа к памяти для каждой итерации гнезда цикла для определенного уровня иерархии памяти:

$$MC(t_1, \dots, t_n) = LC \times \frac{N(t_1, \dots, t_n) \times M(t_1, \dots, t_n)}{t_1 \times \dots \times t_n}, \quad (4)$$

где LC — издержки промаха на строку в кэше и буфере ассоциативной трансляции; $N(t_1, \dots, t_n)$ — функция размеров тайлов; $M(t_1, \dots, t_n)$ — функция формы тайлов; $LC \times N(t_1, \dots, t_n) \times M(t_1, \dots, t_n)$ — издержки переноса всех данных цикла в кэш.

Выбор применяемых преобразований определяется данной моделью и самой возможностью таких преобразований [30]. Многоуровневый тайлинг для локальности может быть далее применен в рамках полиэдральной модели вместо перехода в синтаксическую стадию. Вследствие предполагаемой сложности реализации многоуровневого тайлинга в полиэдральном представлении, решение данной задачи оставлено для дальнейших исследований в рамках настоящей работы.

Вторая стадия метода основана на применении подхода иерархического параметризованного тайлинга для преобразования

кода, представленного в виде абстрактного синтаксического дерева. Иерархический тайлинг увеличивает сложность действий для преобразования тайлов, поскольку каждый уровень разбиения на тайлы имеет свои параметры выбора размера и формы тайлов, а особенности схемы разбиения на тайлы на разных уровнях вступают в конфликт. Это означает, что для минимизации времени выполнения простой выбор формы тайла для каждого уровня в отдельности не приведет к наилучшему выбору. Выбор формы тайла может быть осуществлен для множества уровней с помощью глобального иерархического тайлинга. Таким образом, генерируются размеры тайлов, которые являются символическими параметрами для различных конфигураций иерархической архитектуры памяти.

Эффективный набор тайлов с различными вариантами размеров и формы определяется в соответствии с иерархией кэш-памяти, конфигурацией узлов NUMA и буфера ассоциативной трансляции. Для существующих уровней эвристически определяется степень тайлинга, подходящая для снижения времени прямого доступа к памяти. Модель прямого доступа к памяти определена как

$$T_{i,j} = \alpha \times HL_{i,j} + \beta \times TS_{i,j} + \gamma \times TSH_{i,j}, \quad (5)$$

где $HL_{i,j}$ — уровень иерархии с соответствующими издержками прямого доступа; $TS_{i,j}$ — размер тайла; $TSH_{i,j}$ — форма тайла в том случае, если она рассматривается; α , β , γ — положительные константы. Таким образом, необходимо определить размер и форму тайла для различных уровней иерархии в целях снижения издержек прямого доступа к памяти.

Из формулы, предложенной в работе [38], получаем издержки доступа для каждого уровня иерархической памяти:

$$HL_{i,j} = \sum_{i=1}^N \sum_{j=1}^M \sigma(i,j) j \left(T_{DMA} + \frac{block_i}{BW} \right). \quad (6)$$

Здесь $\sigma(i,j)$ показывает, происходят ли промахи в кэш j раз в i -м блоке данных;

$block_i$ — размер i -го блока данных; T_{DMA} — задержка времени доступа к удаленным данным; BW — пропускная способность памяти при удаленном доступе. С учетом издержек доступа метод предполагает корректировку параметров тайла для различных уровней иерархии $HL_{i,j}$ посредством применения алгоритмов выбора размера тайла (Tile Size Selection — TSS) и формы тайла (Tile Shape Selection — TSHS).

Построение производных алгоритмов.

Предложенный метод является своеобразным шаблоном, допускающим варианты корректировки таких параметров тайла, как размер и форма в зависимости от характеристик доступной иерархической памяти посредством применения подстановочных алгоритмов. Данная возможность предусмотрена в связи с постоянным совершенствованием алгоритмов класса TSS/TSHS. На основе описанного метода были построены варианты 1 и 2 алгоритма аcoгерентного неоднородного тайлинга (Acoherent Non-Uniform Loop Tiling — ACNULT).

Аcoгерентность в данном случае означает, что алгоритм полагается на аппаратную схему обеспечения когерентности кэш-памяти и работает с глобально адресуемой общей памятью терабайтного объема как с памятью обычного вычислительного узла. Варианты ACNULT(1,2) реализуют техники TSS и TSHS соответственно. Подстановочные алгоритмы поиска размеров и формы тайла в рамках алгоритмов ACNULT(1,2) представляют собой поиск по пространству, определяемому эксцентриситетом и обратным рабочим множеством, размер которого зависит от размера кэша. Порядок поиска определяется эвристикой, которая также используется для обрезки пространства поиска. Вопросы построения эффективных алгоритмов типа TSS/TSHS выходят за пределы настоящей статьи и в дальнейшем не рассматриваются. Выводы о степени сложности данных методов можно сделать из работ [31–35]. Общая схема предложенного метода демонстрируется на рис. 1.

```

Вход: Промежуточное представление программы P
Выход: Промежуточное представление оптимизированной программы P

1 begin
2   /* Полиэдральные преобразования с приоритизацией локальности */
3   begin
4     P := (экспериментально) объединение циклов
5     P := перестановка циклов
6     P := (экспериментально) сдвиг цикла/тайлинг
7   end
8   /* Синтаксические преобразования с моделированием неоднородности */
9   begin
10    P := сдвиг цикла
11    P := моделирование прямого доступа к памяти для доступных уровней иерархии
12    P := тайлинг посредством ACNULT(1) с применением метода TSS
13    P := тайлинг посредством ACNULT(2) с применением метода TSHS
14  end
15 end

```

Рис. 1. Метод гибридного неоднородного тайлинга с использованием вариантов ACNULT

Fig. 1. Overall algorithmic skeleton pipeline extended with ACNULT variants

Для реализации предложенного метода используется Polly, инструмент для полиэдральной оптимизации промежуточного представления инфраструктуры компиляторов LLVM [26]. Инструмент Polly применяется для обнаружения в канонизированном коде статических областей управления, которые можно преобразовать в полиэдральное представление и подвергнуть оптимизации. Преобразования в рамках полиэдрального подхода описаны в формате JSCoP, который используется для импорта и экспорта полиэдрального представления с измененными последовательностями выражений.

Экспериментальные результаты

Цель эксперимента — проверка возможностей алгоритмов адекватного неоднородного тайлинга ACNULT в отношении повышения эффективности тестовых полиэдральных программ на машинах с глубокой иерархией памяти.

Экспериментальная система включает набор многомашинных макроузлов с глобально адресуемой когерентной памятью с неоднородным доступом (Cache-Coherent Non-Uniform Memory Access — ccNUMA). Аппаратные конфигурации базовых узлов M0 и варианты построенных на их основе макроузлов M1, M2, M3 представлены в табл. 1. Каждый макроузел представляет со-

бой кластер с общей памятью, при этом максимальные конфигурации макроузлов имеют ресурс процессоров и оперативной памяти, недоступный на узле стандартного вычислительного кластера общего назначения. Так, максимальный узел типа M3 способен предоставить более 12 ТБ RAM одновременно с 3072 ядер процессоров под управлением единственного экземпляра стандартной операционной системы на основе ядра Linux версии 4.14. Поскольку общий объем памяти ограничен только 48-битным диапазоном физических адресов, то общая возможность адресации составляет до 256 ТБ RAM. Глубокая иерархическая структура памяти строится посредством объединения базовых узлов типа M0 в макроузел посредством вычислительной сети с использованием протокола когерентности, основанного на стандарте Scalable Coherent Interface (SCI).

Таким образом, основным архитектурным принципом системы является наличие глобально адресуемого адресного пространства, что и позволяет процессорам осуществлять неоднородный доступ к удаленной памяти непосредственно. Адекватный алгоритм будет рассматривать весь доступный объем памяти как память одного узла. Полный список уровней иерархии памяти экспериментального макроузла представлен в табл. 2.



Таблица 1

Варианты конфигураций экспериментальных макроузлов

Table 1

Target macronode configurations

Особенности архитектуры	Базовый узел (M0)	Варианты конфигурации макроузлов		
		Минимальный (M1)	Средний (M2)	Максимальный (M3)
RAM	188 Гбайт	752 Гбайт	3 ТБ	12 ТБ
Узлы NUMA	6	24	96	384
Плата/Сокет/Ядра	1/3/48	4/12/192	16/48/768	64/192/3072

Таблица 2

Доступные уровни иерархии для моделирования локальности

Table 2

Available levels of deep memories to directly target locality

Уровни	Размер	Определение
L1d cache	16 К	Кэш данных
L1i cache	64 К	Кэш инструкций
L2 cache	2018 К	Кэш второго уровня
L3 cache	6144 К	Общий кэш для ядер, принадлежащих узлу NUMA
NUMA 0	32 Гбайт	Собственный узел NUMA
NUMA 1	32 Гбайт	Соседний узел NUMA по CPU
NUMA 2-5	До 188 Гбайт	NUMA-узлы процессоров внутри одного базового узла
L4 cache	1-16 Гбайт	Конфигурируемый кэш для удаленных данных
NUMA 6-384	До 12 ТБ	NUMA-узлы соседних базовых узлов

Первая стадия экспериментальной проверки возможностей алгоритмов ACNULT содержит оценку эффективности преобразованного кода для задач класса трафаретных вычислений из набора тестовых программ PolyBench/C, который обычно используется для оценки полиэдральных оптимизаторов [4, 26, 27]. Избранные задачи данного класса включают неявный метод переменного направления для двумерной диффузии тепла *adi*, метод конечных разностей во временной области для двумерных данных *fdtd-2d*, решение уравнения теплопроводности на трёхмерном пространстве *heat-3d*. Кроме того, используются трафаретные вычисления методом

Якоби на одномерных или двумерных данных с использованием трёхточечного или пятиточечного шаблона *jacobi-1d* и *jacobi-2d*, а также вычисления методом Гаусса-Зейделя по двумерным данным с использованием шаблона с девятью точками *seidel-2d*. Данные программы использовались с максимальными настройками размеров данных *LARGE* и *EXTRALARGE* для создания репрезентативных настроек на узле типа M0. Иерархия памяти включает уровни кэшей L1d/L1i/L2/L3 наряду с трёхуровневой иерархией узлов NUMA и с конфигурируемой памятью L4 для обеспечения работы протокола когерентности кэшей.

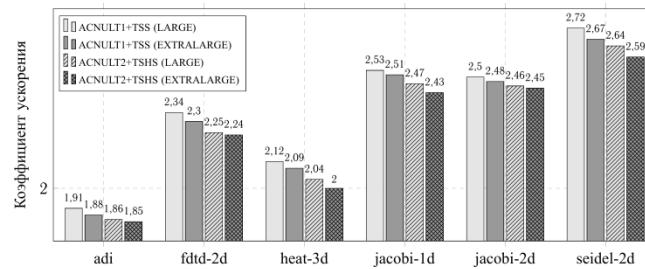


Рис. 2. Ускорение трафаретных вычислений PolyBench/C при использовании вариантов ACNULT

Fig. 2. Speedup with ACNULT on PolyBench/C stencils

На рис. 2 результаты алгоритмов ACNULT(1,2) сопоставляются с результатами кода, полученного только с параметром компилятора clang -O3 без использования Polly. Улучшения при использовании Polly здесь обусловлены применением тайлинга и выбором алгоритмов TSS/TSHS для подстановки в алгоритм ACNULT. Для случая более крупных узлов ускорение будет определяться в основном посредством выбора алгоритмов TSS. Результаты EXTRALARGE близки к результатам LARGE, что объясняется использованием тестовыми задачами недостаточных размеров данных в случае максимальной настройки EXTRALARGE, в результате чего демонстрируется некоторое ускорение.

Вторая стадия экспериментальной проверки метода включает исследование локальности как фактора переносимости производительности. Проблема переносимости производительности состоит в недостаточной эффективности вычислительного алгоритма при компиляции и выполнении на архитектурах, созданных с использованием принципов, отличных от исходных подходов. Для оценки переносимости производительности целесообразно использовать прокси-программы, которые обычно применяются при суперкомпьютерном кодизаине микроархитектуры и компилятора [39]. Переносимость производительности оптимизированной прокси-программы может быть определена в соответствии с моделью, ранее предложенной в [14], и интерпретированной в [9] специально для случая множества конфигураций

макроузлов $|M|$, каждая последующая из которых содержит в себе предыдущие и, таким образом, иерархия памяти углубляется по мере их укрупнения:

$$PP_{m-node, opt, proxy} = \begin{cases} \frac{|M|}{\sum_{i \in M} \frac{1}{T(opt, proxy)}}, & \text{if } \forall i \in M, \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

где $T(opt, proxy)$ — время выполнения оптимизированной прокси-программы для макроузла i .

Переносимость производительности будет достигнута лишь в том случае, когда улучшение производительности оптимизированной прокси-программы будет зафиксировано на всех целевых машинах. В данной модели локальность может быть рассмотрена как основной фактор, влияющий на переносимость производительности. Метод оценки локальности рассмотрен в работе [9]. Для оценки улучшения локальности применяется программа Livermore Unstructured Lagrange Explicit Shock Hydrodynamics (LULESH), представляющая собой суррогатный код, репрезентативный широкому спектру научных вычислительных задач в области гидродинамики, выполняемых на суперкомпьютерах. LULESH является научной прокси-программой, что предполагает наличие ряда алгоритмически эквивалентных реализаций данной программы с использованием множества перспективных моделей параллельных вычислений (MPI/OpenMP, Charm++, Chapel, MPI/UPC++, X10, etc.), а также наличие результатов оценки произво-

дительности на таких целевых архитектурах суперкомпьютеров, как BlueGene/Q, Cray XE6, POWER9, AMC [39].

Формула (8) ассоциирует вычислительные прокси-процедуры, из которых состоит прокси-программа, с набором преобразований кода [9]. Время выполнения прокси-программы определяется как

$$T = \sum_{i=1}^n Kernel_{i,opt,proxy} T_{i,m-node,opt}, \quad (8)$$

где $Kernel_{i,opt,proxy}$ — число оптимизируемых прокси-процедур; $T_{i,m-node,opt}$ — время выполнения прокси-процедуры i на макроузле после применения преобразования opt .

Код LULESH можно рассматривать как в качестве прокси-программы, так и в качестве полиэдрального теста производительности для суперкомпьютеров. Для этого используется реализация LULESH, основанная на использовании модели параллельных вычислений OpenMP, что является следствием архитектурных особенностей макроузлов, тяготеющих к укрупнению (от конфигурации M1 до M3). При этом реализация OpenMP наряду с версией MPI наиболее часто применяется в области высокопроизводительных вычислений и поэтому имеет множество опубликованных результатов для других высокопроизводительных проблемно-ориентированных архитектур, что позволяет сопоставить эти системы.

Запуск более тысячи потоков OpenMP внутри одного узла типа M3 представляет собой стресс-тест для многомашинной системы с общей памятью [9]. Однако, в отличие от тестов PolyBench/C, изначально разработанных для определения SCoP, LULESH содержит более сложные SCoP, которые могут быть сформированы посредством преобразования из наиболее трудоемких параллельных регионов OpenMP. Таким образом, полученные статические области управления могут быть преобразованы в полиэдральное представление при условии, что множественные избыточные зависимости между различными операторами, содержащиеся в SCoP, будут устранены.

Результаты оценки пространственной локальности LULESH после применения алгоритмов ACNULT(1,2) относительно локальности первоначальной версии представлены на рис. 3. Многопоточная реализация LULESH использована с размером задачи $30^3 \dots 90^3$ на макроузлах с 752 Гбайт, 3 ТБ, 12 ТБ глобально адресуемой памяти. Видно, что расслоение поверхностей наблюдается по мере углубления иерархической памяти от узла 3 ТБ и далее для всех размеров задач. Вариант ACNULT1 использует технику TSS. Данный вариант представляется более эффективным по сравнению с ACNULT2+TSHS. Улучшение локальности сохраняется по мере углубления памяти макроузлов для двух вариантов алгоритма. Кроме того, были улучшены характеристики локальности LULESH при увеличении размера задачи до значения 90^3 сложности. В результате при всех размерах задачи LULESH локальность при использовании ACNULT(1,2) оказалась лучше локальности неоптимизированного кода при использовании наиболее крупномасштабного макроузла M3 с наибольшей степенью неоднородности памяти, что, согласно формуле (7), свидетельствует о достижении переносимости производительности на все целевые машины.

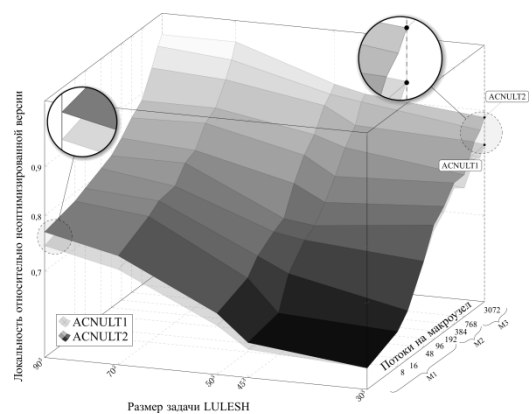


Рис. 3. Локальность многопоточной реализации LULESH для задачи $30^3 \dots 90^3$ на макроузлах с 752 Гбайт, 3 ТБ, 12 ТБ RAM

Fig. 3. Locality results for multithreaded LULESH with $30^3 \dots 90^3$ sized mesh on the macronodes with 752 Gb, 3 Tb, 12 Tb RAM

Закключение

В статье предложен метод гибридного неоднородного тайлинга для архитектур суперкомпьютеров с многоуровневой иерархией памяти. Данный метод основан на комбинировании полиэдрального и синтаксического подходов автоматического преобразования циклов. Разработана последовательность преобразований, преимущества которой перед существующими подходами состоят в возможности осуществить приоритизацию локальности на полиэдральной стадии и моделировать аспекты неоднородности целевой архитектуры на стадии синтаксического иерархического тайлинга.

Построены алгоритмы акогогерентного неоднородного тайлинга, позволяющие расширить область применения предлагаемого метода посредством использования перспективных подстановочных техник для тайлинга. Получены экспериментальные результаты построенных алгоритмов для повышения эффективности трафаретных вычислений и для улучшения локальности полиэдральной прокси-программы на макроузлах с глобально адресуемой когерентной памятью с неоднородным доступом.

Дальнейшие исследования будут направлены на совершенствование методов многоуровневого тайлинга в рамках полиэдральной модели компиляции.

СПИСОК ЛИТЕРАТУРЫ

1. Benabderrahmane M., Pouchet L., Cohen A., et al. The polyhedral model is more widely applicable than you think // Proc. of the 19th Joint European Conf. on Theory and Practice of Software, Internat. Conf. on Compiler Construction. 2010. Pp. 283–303. DOI: 10.1007/978-3-642-11970-5_16
2. Hammami E., Slama Y. An overview on loop tiling techniques for code generation // 2017 IEEE/ACS 14th Internat. Conf. on Computer Systems and Applications. 2017. Pp. 280–287. DOI: 10.1109/AICCSA.2017.168
3. Wijesinghe T., Senevirathne K., Siriwardhana C., et al. Parameterized diamond tiling for parallelizing stencil computations // Moratuwa Engineering Research Conf. 2017. Pp. 99–104. DOI: 10.1109/MERCON.2017.7980464
4. Zinenko O., Verdoolaege S., Reddy C., et al. Modeling the conflicting demands of parallelism and temporal/spatial locality in affine scheduling // Proc. of the 27th Internat. Conf. on Compiler Construction. 2018. Pp. 3–13. DOI: 10.1145/3178372.3179507
5. Zinenko O., Chelini L., Grosser T. Declarative transformations in the polyhedral model. Research report RR-9243. Inria, ENS Paris – Ecole Normale Supérieure de Paris, ETH Zurich, TU Delft, IBM Zurich, 2018 // URL: <https://hal.inria.fr/hal-01965599> (Дата обращения: 30.04.2019).
6. Zinenko O., Huot S., Bastoul C. Visual program manipulation in the polyhedral model // ACM Transactions on Architecture and Code Optimization. 2018. Vol. 15. No. 1. Pp. 1–25. DOI: 10.1145/3177961
7. Li F., Qiu K., Zhao M., et al. Checkpointing-aware loop tiling for energy harvesting powered nonvolatile processors // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2019. Vol. 38. No. 1. Pp. 15–28. DOI: 10.1109/TCAD.2018.2803624
8. Tanase A., Hannig F., Teich J. Symbolic parallelization. Symbolic parallelization of nested loop programs. 2018. Pp. 37–92. DOI: 10.1007/978-3-319-73909-0_3
9. Levchenko A. Exploring trade-offs of compiler optimizations to enable performance portability for multi-level memory hierarchies // Proc. of the 4rd Ural Workshop on Parallel, Distributed, and Cloud Computing for Young Scientists. 2018. Pp. 85–97. Available: <http://ceur-ws.org/Vol-2281/> (Accessed: 30.04.2019)
10. Zhai J., Chen W. A vision of post-exascale programming // Frontiers of Information Technology and Electronic Engineering. 2018. Vol. 19. No. 10. Pp. 1261–1266. DOI: 10.1631/FITEE.1800442
11. Yang G., Fu H. Application software beyond exascale: challenges and possible trends // Frontiers of Information Technology and Electronic Engineering. 2018. Vol. 19. No. 10. Pp. 1267–1272. DOI: 10.1631/FITEE.1800459
12. Chen Z., Dongarra J., Xu Z. Post-exascale supercomputing: research opportunities abound // Frontiers of Information Technology and Electronic Engineering. 2018. Vol. 19. No. 10. Pp. 1203–1208. DOI: 10.1631/FITEE.1830000
13. Eisymont L. Hybrid strategy development of the supercomputer components // Open Systems. DBMS. 2017. Vol. 25. No. 2. Pp. 8–11.
14. Pennycook S., Sewall J., Lee V. Implications of a metric for performance portability // Future



- Generation Computer Systems. 2019. Vol. 92. Pp. 947–958. DOI: 10.1016/J.FUTURE.2017.08.007
15. **Zounmevo J., Perarnau S., Iskra K., et al.** A container-based approach to OS specialization for exascale computing // 2015 IEEE Internat. Conf. on Cloud Engineering. 2015. Pp. 359–364. DOI: 10.1109/IC2E.2015.78
16. **Ashraf R., Gioiosa R., Kestor G., et al.** Exploring the effect of compiler optimizations on the reliability of HPC applications // 2017 IEEE International Parallel and Distributed Processing Symp. Workshops. 2017. Pp. 1274–1283. DOI: 10.1109/IPDPSW.2017.7
17. **Perarnau S., Zounmevo J., Dreher M., et al.** Argo NodeOS: toward unified resource management for exascale // 2017 IEEE Internat. Parallel and Distributed Processing Symp. 2017. Pp. 153–162. DOI: 10.1109/IPDPS.2017.25
18. **Zou P., Allen T., Davis C., et al.** CLIP: cluster-level intelligent power coordination for power-bounded systems // 2017 IEEE Internat. Conf. on Cluster Computing. 2017. Pp. 541–551. DOI: 10.1109/CLUSTER.2017.98
19. **Goglin B.** Memory footprint of locality information on many-core platforms // 2018 IEEE Internat. Parallel and Distributed Processing Symp. Workshops. 2018. Pp. 1283–1292. DOI: 10.1109/IPDPSW.2018.00201
20. **Rodriguez G., Kandemir M., Tourino J.** Affine modeling of program traces // IEEE Transactions on Computers. 2019. Vol. 68. No. 2. Pp. 294–300. DOI: 10.1109/TC.2018.2853747
21. **Siklosi B., Reguly I., Mudalige G.** Heterogeneous CPU-GPU execution of stencil applications // 2018 IEEE/ACM Internat. Workshop on Performance, Portability and Productivity in HPC. 2018. Pp. 71–80. DOI: 10.1109/P3HPC.2018.00010
22. **Simburger A., Apel S., Grosslinger A., et al.** PolyJIT: polyhedral optimization just in time // Internat. J. of Parallel Programming, 2018. DOI: 10.1007/S10766-018-0597-3
23. **Bielecki W., Palkowski M., Skotnicki P.** Generation of parallel synchronization-free tiled code // Computing. 2018. Vol. 100. No. 3. Pp. 277–302. DOI: 10.1007/S00607-017-0576-3
24. **Liu S., Cui Y., Zou N., et al.** Revisiting the parallel strategy for DOACROSS loops // J. of Computer Science and Technology. 2019. Vol. 34. No. 2. Pp. 456–475. DOI: 10.1007/S11390-019-1919-7
25. **Zinenko O., Verdoolaege S., Reddy C., et al.** Unified polyhedral modeling of temporal and spatial locality. Research report RR-9110. Inria Paris, 2017 // URL: <https://hal.inria.fr/hal-01628798> (Дата обращения: 30.04.2019).
26. **Grosser T., Grosslinger A., Lengauer C.** Polly – performing polyhedral optimizations on a low-level intermediate representation // Parallel Processing Letters. 2012. Vol. 22. No. 04. DOI: 10.1142/S0129626412500107
27. **Kruse M., Grosser T.** DeLICM: scalar dependence removal at zero memory cost // Proc. of the 2018 Internat. Symp. on Code Generation and Optimization. 2018. Pp. 241–253. DOI: 10.1145/3168815
28. **Acharya A., Bondhugula U., Cohen A.** Polyhedral auto-transformation with no integer linear programming // Proc. of the 39th ACM SIGPLAN Conf. on Programming Language Design and Implementation. 2018. Vol. 53. No. 4. Pp. 529–542. DOI: 10.1145/3192366.3192401
29. **Pradelle B., Meister B., Baskaran M., et al.** Scalable hierarchical polyhedral compilation // 45th Internat. Conf. on Parallel Processing. 2016. Pp. 432–441. DOI: 10.1109/ICPP.2016.56
30. **Shirako J., Pouchet L., Sarkar V.** Oil and water can mix: an integration of polyhedral and AST-based transformations // SC14: Proc. of the Internat. Conf. for High Performance Computing, Networking, Storage and Analysis. 2014. Pp. 287–298. DOI: 10.1109/SC.2014.29
31. **Liu J., Wickerson J., Constantinides G.** Tile size selection for optimized memory reuse in high-level synthesis // 27th Internat. Conf. on Field Programmable Logic and Applications. 2017. Pp. 1–8. DOI: 10.23919/FPL.2017.8056810
32. **Chen Y., Shang W., Fang Y.** Supernode transformation on computer clusters // 10th Internat. Conf. on Ubi-media Computing and Workshops. 2017. Pp. 1–6. DOI: 10.1109/UMEDIA.2017.8074109
33. **Bondhugula U., Bandishti V., Pananilath I.** Diamond Tiling: tiling techniques to maximize parallelism for stencil computations // IEEE Transactions on Parallel and Distributed Systems. 2017. Vol. 28. No. 5. Pp. 1285–1298. DOI: 10.1109/TPDS.2016.2615094
34. **Reguly I., Mudalige G., Giles M.** Loop tiling in large-scale stencil codes at run-time with OPS // IEEE Transactions on Parallel and Distributed Systems. 2018. Vol. 29. No. 4. Pp. 873–886. DOI: 10.1109/TPDS.2017.2778161
35. **Rawat P., Vaidya M., Sukumaran-Rajam A., et al.** Domain-specific optimization and generation of high-performance GPU code for stencil computations // Proc. of the IEEE. 2018. Vol. 106. No. 11. Pp. 1902–1920. DOI: 10.1109/JPROC.2018.2862896
36. **Qiu K., Ni Y., Zhang W., et al.** An adaptive non-uniform loop tiling for DMA-based bulk data transfers on many-core processor // 2016 IEEE

34th Internat. Conf. on Computer Design. 2016. Pp. 9–16. DOI: 10.1109/ICCD.2016.7753255

37. **Razanajato H., Bastoul C., Loechner V.** Lifting barriers using parallel polyhedral regions // IEEE 24th Internat. Conf. on High Performance Computing. 2017. Pp. 338–347. DOI: 10.1109/HIPC.2017.00046

38. **Zhang M., Gu N., Ren K.** Optimization of computation-intensive applications in cc-NUMA

architecture // Internat. Conf. on Networking and Network Applications. 2016. Pp. 244–249. DOI: 10.1109/NaNA.2016.12

39. **Jacob A., Nair R., Chen T., et al.** Progressive codesign of an architecture and compiler using a proxy application // 27th Internat. Symp. on Computer Architecture and High Performance Computing. 2015. Pp. 57–64. DOI: 10.1109/SBAC-PAD.2015.18

Статья поступила в редакцию 30.04.2019.

REFERENCES

1. **Benabderrahmane M., Pouchet L., Cohen A., et al.** The polyhedral model is more widely applicable than you think. *Proceedings of the 19th Joint European Conference on Theory and Practice of Software, International Conference on Compiler Construction*, 2010, Pp. 283–303. DOI: 10.1007/978-3-642-11970-5_16

2. **Hammami E., Slama Y.** An overview on loop tiling techniques for code generation. *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications*, 2017, Pp. 280–287. DOI: 10.1109/AICCSA.2017.168

3. **Wijesinghe T., Senevirathne K., Siriwardhana C., et al.** Parameterized diamond tiling for parallelizing stencil computations. *Moratuwa Engineering Research Conference*, 2017, Pp. 99–104. DOI: 10.1109/MERCON.2017.7980464

4. **Zinenko O., Verdoolaege S., Reddy C., et al.** Modeling the conflicting demands of parallelism and Temporal/Spatial locality in affine scheduling. *Proceedings of the 27th International Conference on Compiler Construction*, 2018, Pp. 3–13. DOI: 10.1145/3178372.3179507

5. **Zinenko O., Chellini L., Grosser T.** Declarative transformations in the polyhedral model. Research report RR-9243. Inria, ENS Paris – Ecole Normale Supérieure de Paris, ETH Zurich, TU Delft, IBM Zurich, 2018. Available: <https://hal.inria.fr/hal-01965599> (Accessed: 30.04.2019).

6. **Zinenko O., Huot S., Bastoul C.** Visual program manipulation in the polyhedral model. *ACM Transactions on Architecture and Code Optimization*, 2018, Vol. 15, No. 1, Pp. 1–25. DOI: 10.1145/3177961

7. **Li F., Qiu K., Zhao M., et al.** Checkpointing-aware loop tiling for energy harvesting powered nonvolatile processors. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2019, Vol. 38, No. 1, Pp. 15–28. DOI: 10.1109/TCAD.2018.2803624

8. **Tanase A., Hannig F., Teich J.** Symbolic parallelization. *Symbolic parallelization of nested loop programs*, 2018, Pp. 37–92. DOI: 10.1007/978-3-319-73909-0_3

9. **Levchenko A.** Exploring trade-offs of compiler optimizations to enable performance portability for multi-level memory hierarchies. *Proceedings of the 4rd Ural Workshop on Parallel, Distributed, and Cloud Computing for Young Scientists*, 2018, Pp. 85–97. Available: <http://ceur-ws.org/Vol-2281/> (Accessed: 30.04.2019).

10. **Zhai J., Chen W.** A vision of post-exascale programming. *Frontiers of Information Technology and Electronic Engineering*, 2018, Vol. 19, No. 10, Pp. 1261–1266. DOI: 10.1631/FITEE.1800442

11. **Yang G., Fu H.** Application software beyond exascale: challenges and possible trends. *Frontiers of Information Technology and Electronic Engineering*, 2018, Vol. 19, No. 10, Pp. 1267–1272. DOI: 10.1631/FITEE.1800459

12. **Chen Z., Dongarra J., Xu Z.** Post-exascale supercomputing: research opportunities abound. *Frontiers of Information Technology and Electronic Engineering*, 2018, Vol. 19, No. 10, Pp. 1203–1208. DOI: 10.1631/FITEE.1830000

13. **Eisymont L.** Hybrid strategy development of the supercomputer components. *Open Systems. DBMS*, 2017, Vol. 25, No. 2, Pp. 8–11.

14. **Pennycook S., Sewall J., Lee V.** Implications of a metric for performance portability. *Future Generation Computer Systems*, 2019, Vol. 92, Pp. 947–958. DOI: 10.1016/J.FUTURE.2017.08.007

15. **Zounmevo J., Perarnau S., Iskra K., et al.** A container-based approach to OS specialization for exascale computing. *2015 IEEE International Conference on Cloud Engineering*, 2015, Pp. 359–364. DOI: 10.1109/IC2E.2015.78

16. **Ashraf R., Gioiosa R., Kestor G., et al.** Exploring the effect of compiler optimizations on the reliability of HPC applications. *2017 IEEE*



International Parallel and Distributed Processing Symposium Workshops, 2017, Pp. 1274–1283. DOI: 10.1109/IPDPSW.2017.7

17. **Perarnau S., Zounmevo J., Dreher M., et al.** Argo NodeOS: toward unified resource management for exascale. *2017 IEEE International Parallel and Distributed Processing Symposium*, 2017, Pp. 153–162. DOI: 10.1109/IPDPS.2017.25

18. **Zou P., Allen T., Davis C., et al.** CLIP: cluster-level intelligent power coordination for power-bounded systems. *2017 IEEE International Conference on Cluster Computing*, 2017, Pp. 541–551. DOI: 10.1109/CLUSTER.2017.98

19. **Goglin B.** Memory footprint of locality information on many-core platforms. *2018 IEEE International Parallel and Distributed Processing Symposium Workshops*, 2018, Pp. 1283–1292. DOI: 10.1109/IPDPSW.2018.00201

20. **Rodriguez G., Kandemir M., Tourino J.** Affine modeling of program traces. *IEEE Transactions on Computers*, 2019, Vol. 68, No. 2, Pp. 294–300. DOI: 10.1109/TC.2018.2853747

21. **Siklosi B., Reguly I., Mudalige G.** Heterogeneous CPU-GPU execution of stencil applications. *2018 IEEE/ACM International Workshop on Performance, Portability and Productivity in HPC*, 2018, Pp. 71–80. DOI: 10.1109/P3HPC.2018.00010

22. **Simburger A., Apel S., Grosslinger A., et al.** PolyJIT: polyhedral optimization just in time. *International Journal of Parallel Programming*, 2018. DOI: 10.1007/S10766-018-0597-3

23. **Bielecki W., Palkowski M., Skotnicki P.** Generation of parallel synchronization-free tiled code. *Computing*, 2018, Vol. 100, No. 3, Pp. 277–302. DOI: 10.1007/S00607-017-0576-3

24. **Liu S., Cui Y., Zou N., et al.** Revisiting the parallel strategy for DOACROSS loops. *Journal of Computer Science and Technology*, 2019, Vol. 34, No. 2, Pp. 456–475. DOI: 10.1007/S11390-019-1919-7

25. **Zinenko O., Verdoolaege S., Reddy C., et al.** Unified polyhedral modeling of temporal and spatial locality. Research report RR-9110. Inria Paris, 2017. Available: <https://hal.inria.fr/hal-01628798> (Accessed: 30.04.2019).

26. **Grosser T., Grosslinger A., Lengauer C.** Polly – performing polyhedral optimizations on a low-level intermediate representation. *Parallel Processing Letters*, 2012, Vol. 22, No. 04. DOI: 10.1142/S0129626412500107

27. **Kruse M., Grosser T.** DeLICM: scalar dependence removal at zero memory cost. *Proceedings of the 2018 International Symposium on*

Code Generation and Optimization, 2018, Pp. 241–253. DOI: 10.1145/3168815

28. **Acharya A., Bondhugula U., Cohen A.** Polyhedral auto-transformation with no integer linear programming. *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design and Implementation*, 2018, Vol. 53, No. 4, Pp. 529–542. DOI: 10.1145/3192366.3192401

29. **Pradelle B., Meister B., Baskaran M., et al.** Scalable hierarchical polyhedral compilation. *2016 45th International Conference on Parallel Processing*, 2016, Pp. 432–441. DOI: 10.1109/ICPP.2016.56

30. **Shirako J., Pouchet L., Sarkar V.** Oil and water can mix: an integration of polyhedral and AST-based transformations. *SC14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2014, Pp. 287–298. DOI: 10.1109/SC.2014.29

31. **Liu J., Wickerson J., Constantinides G.** Tile size selection for optimized memory reuse in high-level synthesis. *2017 27th International Conference on Field Programmable Logic and Applications*, 2017, Pp. 1–8. DOI: 10.23919/FPL.2017.8056810

32. **Chen Y., Shang W., Fang Y.** Supernode transformation on computer clusters. *2017 10th International Conference on Ubi-media Computing and Workshops*, 2017, Pp. 1–6. DOI: 10.1109/UMEDIA.2017.8074109

33. **Bondhugula U., Bandishti V., Pananilath I.** Diamond tiling: tiling techniques to maximize parallelism for stencil computations. *IEEE Transactions on Parallel and Distributed Systems*, 2017, Vol. 28, No. 5, Pp. 1285–1298. DOI: 10.1109/TPDS.2016.2615094

34. **Reguly I., Mudalige G., Giles M.** Loop tiling in large-scale stencil codes at run-time with OPS. *IEEE Transactions on Parallel and Distributed Systems*, 2018, Vol. 29, No. 4, Pp. 873–886. DOI: 10.1109/TPDS.2017.2778161

35. **Rawat P., Vaidya M., Sukumaran-Rajam A., et al.** Domain-specific optimization and generation of high-performance GPU code for stencil computations. *Proceedings of the IEEE*, 2018, Vol. 106, No. 11, Pp. 1902–1920. DOI: 10.1109/JPROC.2018.2862896

36. **Qiu K., Ni Y., Zhang W., et al.** An adaptive non-uniform loop tiling for DMA-based bulk data transfers on many-core processor. *2016 IEEE 34th International Conference on Computer Design*, 2016, Pp. 9–16. DOI: 10.1109/ICCD.2016.7753255

37. **Razanajato H., Bastoul C., Loechner V.** Lifting barriers using parallel polyhedral regions. *2017 IEEE 24th International Conference on High Performance Computing*, 2017, Pp. 338–347. DOI: 10.1109/HIPC.2017.00046

38. Zhang M., Gu N., Ren K. Optimization of computation-intensive applications in cc-NUMA architecture. *2016 International Conference on Networking and Network Applications*, 2016, Pp. 244–249. DOI: 10.1109/NaNA.2016.12

39. Jacob A., Nair R., Chen T., et al. Progressive codesign of an architecture and compiler using a proxy application. *2015 27th International Symposium on Computer Architecture and High Performance Computing*, 2015, Pp. 57–64. DOI: 10.1109/SBAC-PAD.2015.18

Received 30.04.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

ЛЕВЧЕНКО Алексей Викторович

LEVCHENKO Aleksei V.

E-mail: levchenko_av@spbstu.ru

DOI: 10.18721/JCSTCS.12404
УДК 004.896

ПРОГРАММНЫЙ КОМПЛЕКС УПРАВЛЕНИЯ МАТЕРИАЛЬНЫМИ РЕСУРСАМИ ЛИСТОВОГО МЕТАЛЛА С ЭЛЕМЕНТАМИ СППР

А.В. Улитин¹, А.А. Смирнов²

¹ Вологодский государственный университет,
Вологда, Российская Федерация;

² Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

На машиностроительных предприятиях, где существенную долю материальных затрат занимают затраты на листовой металл, рациональное использование материальных ресурсов листового металла влияет на технологические затраты и на другие технические и экономические показатели операционной деятельности предприятия. Именно поэтому на машиностроительных предприятиях актуален вопрос увеличения коэффициента использования листового металла. В целях решения данного вопроса в статье изучено применение методов рационального раскроя листового металла и использование остатков листового металла после раскроя в последующих технологических процессах. Задача данного исследования — разработка методики сортировки остатков раскроя листового металла и построение программного комплекса управления материальными ресурсами листового металла на базе предложенной методики. Разработанный программный комплекс включает в себя систему поддержки принятия решений. Предлагаемая СППР помогает в оперативном формировании рациональных карт раскроя и в оперативной и обоснованной сортировке остатков после раскроя по группам деловых и неделовых остатков. Разработанный программный комплекс позволяет увеличить коэффициент использования листового металла. В статье перечислены функции программного комплекса, описана его структура, сформулированы особенности его построения на базе взаимодействующих CAD/CAM и ERP-систем, предложены алгоритмы работы и практическая реализация отдельных программных модулей.

Ключевые слова: раскрой листового металла, коэффициент использования материала, деловые остатки листового металла, СППР, CAD, CAM, ERP.

Ссылка при цитировании: Улитин А.В., Смирнов А.А. Программный комплекс управления материальными ресурсами листового металла с элементами СППР // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 45–57. DOI: 10.18721/JCSTCS.12404

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

SOFTWARE PACKAGE FOR MANAGING MATERIAL RESOURCES OF SHEET METAL WITH ELEMENTS OF DSS

A.V. Ulitin¹, A.A. Smirnov²

¹ Vologda State University,
Vologda, Russian Federation;

² Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

At engineering enterprises, where a significant share of material costs are the sheet metal costs, the rational use of the sheet metal material resources has an influence on the technological costs and other technical and economic indicators of the enterprise

operating activities. That is why the question of increasing a sheet metal utilization rate is relevant at the engineering enterprises. In order to solve this question, the using of rational cutting methods of sheet metal and the using of sheet metal residues after cutting in following technological processes are researched in the work. The objective of the research is to develop a software package, which includes a decision support system, for managing of sheet metal material resources. The DSS will help in operational making of rational cutting cards and in operational and substantiated sorting of the cutting residues into business and non-business ones. The software package will allow increasing the sheet metal utilization rate. Requirements for the software are presented, its structure is described, the features of its construction based on interacting CAD/CAM and ERP-systems are formulated, the algorithms of individual software modules and these practical implementation are proposed in the article.

Keywords: sheet metal cutting, material utilization rate, business residues of sheet metal, DSS, CAD, CAM, ERP.

Citation: Ulitin A.V., Smirnov A.A. Software package for managing material resources of sheet metal with elements of DSS. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 45–57. DOI: 10.18721/JCSTCS.12404

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Одной из основных технологических операций на машиностроительных предприятиях является раскрой листового металла на определенные заготовки. При этом использование остатков материала после раскрой влияет на значение технологических затрат производимой продукции, а также на другие экономические и технические показатели операционной деятельности предприятия. Поэтому задача увеличения коэффициента использования листового металла актуальна для машиностроительных предприятий.

Коэффициент использования листового металла определяется как численное отношение массы металла в готовом изделии к массе металла, израсходованного на производство этого изделия. Задача его увеличения может решаться применением методов рационального раскрой исходного листового материала и использованием остатков листового металла после раскрой в последующих технологических процессах. Рациональное управление деловыми остатками в совокупности с освоением методов рационального раскрой имеет большое значение и в области ресурсосберегающих технологий, поскольку его применение

напрямую ведет к экономии исходного материала и к сокращению отходов технологических процессов.

Постановка и описание задачи исследования

Под раскрой листового металла понимают принятое расположение контуров заготовок на исходном листе материала. При этом для повышения коэффициента использования металла лица, ответственные за раскрой, стараются организовать раскрой таким образом, чтобы он обеспечивал минимальный расход исходного материала.

Математически задача рационального раскрой является NP-трудной задачей, под которой понимается наиболее выгодное в определенных условиях расположение геометрических объектов в заданных областях.

Задачу раскрой как задачу оптимизации можно определить следующим образом. Пусть $A_1, A_2, \dots, A_n$ — это заготовки, а $B_1, B_2, \dots, B_m$ — области их размещения. При этом расположение каждой конкретной заготовки A_i определяется тремя координатами x_i, y_i, ϕ_i , где x_i, y_i — это абсцисса и ордината фиксированной точки, а ϕ_i — угол поворота заготовки на плоскости. Тогда задача раскрой сводится к определению параметров размещения заготовок, при ко-

торых некоторая целевая функция $F(x_1, y_1, \varphi_1, x_2, y_2, \varphi_2, \dots, x_n, y_n, \varphi_n)$ достигает своего экстремума. При этом должны выполняться условия размещения заготовок внутри одной из областей размещения $B_1, B_2, \dots, B_m$, условия взаимного непересечения заготовок и условия, определяемые технологическими особенностями процесса раскроя, т. е.

$$F(x_1, y_1, \varphi_1, x_2, y_2, \varphi_2, \dots, x_n, y_n, \varphi_n) \rightarrow \text{extr}, \quad (1)$$

$$f1_{ij}(x_i, y_i, \varphi_i, x_j, y_j, \varphi_j) \geq 0; \quad (2)$$

$$i \neq j; \quad i, j = 1, 2, \dots, n,$$

$$f2_i(x_i, y_i, \varphi_i) \geq 0; \quad i = 1, 2, \dots, n, \quad (3)$$

$$f3_l(x_1, y_1, \varphi_1, x_2, y_2, \varphi_2, \dots, x_n, y_n, \varphi_n) \geq 0; \quad (4)$$

$$l = 1, 2, \dots, L,$$

где (2) — условия взаимного непересечения объектов, (3) — условия расположения в области размещения, (4) — прочие условия. В нашем случае целевой функцией F является функция, значение которой равно коэффициенту использования металла k :

$$k = \frac{\sum_{i=1}^n S_i}{P}, \quad (5)$$

где S_i — площадь i -й заготовки; P — суммарная площадь использованного листа металла.

Ограничения (2)–(4) в аналитическом виде неизвестны, по этой причине для решения задачи раскроя не существуют аналитические методы решения. При решении такого рода задач используются в основном приближенные вычислительные методы. Точные методы в данном случае применимы лишь к задачам с небольшой размерностью n и с сильными ограничениями на геометрию заготовок.

Существует множество теоретических трудов и практических разработок, посвященных применению алгоритмов рационального раскроя промышленных материалов. Первыми отечественными работами, положившими начало исследованиям в области методов рационального раскроя, были работы Л.В. Канторовича и В.А. Залгаллера. На заре научных поисков в области

рационального раскроя многие предлагаемые способы оптимизации были основаны на линейных моделях и методах линейного программирования [1]. Позже были разработаны точные методы, с использованием методологии «ветвей и границ», описанные в работах И.В. Романовского. Сеточный метод генерирования раскроев с максимальной оценкой предложен Э.А. Мухачевой. В настоящее время для решения задач раскроя чаще применяются простые однопроходные и многопроходные эвристики, а также метаэвристические алгоритмы, такие как генетический алгоритм или алгоритм поиска с запретами. Исследованием задачи раскроя-упаковки занимается научная группа ESICUP — Euro Special Interest Group on Cutting and Packing (Европейская специальная группа по задачам в области раскроя упаковки), которая проводит ежегодные международные конференции [2–6].

Поскольку по вопросам рационального раскроя исследователями накоплен большой теоретический материал, то это привело к тому, что на сегодняшний день рынок информационных технологий предлагает удобное программное обеспечение (ПО) с широким функционалом, способствующее оптимизации фигурного раскроя с помощью различных математических алгоритмов. Такого рода ПО предназначено для создания карт раскроя и вывода управляющих программ. Оно относится к классу CAD/CAM-систем, позволяющих проектировать на ЭВМ карты раскроя и формировать управляющие программы для оборудования с ЧПУ. Примерами таких программных средств являются: Техтран, Интех-Раскрой, Сириус, Астра Раскрой, Т-Flex, Lantek Expert, Wrykrys, Aptia Solutions, ProNest, Mazak Smart System и др. Многие из этих программных продуктов позволяют сохранить остаток раскроя, но ни один из них не оказывает помощи специалисту, ответственному за раскрой, в принятии решения по отнесению остатка раскроя в группу деловых или неделовых остатков, т. е. не реализуют СППР. Актуальность данного вопроса и предложения по организаци-

онно-экономическим особенностям его решения отмечены в статьях [7, 8]. Актуальность разработки СППР также подтверждена работами [9–15].

При производстве продукции из деловых остатков целесообразно определенным образом организовать процессы маркировки, хранения, учета и поиска остатков. По этой причине для выработки решения по отнесению остатка раскроя в группу деловых или неделовых материальных ресурсов необходимо взаимодействие применяемой CAD/CAM-системы и учетной системы предприятия, которая относится к классу ERP-систем. ERP-системы располагают достаточной информацией о ресурсах предприятия для формирования решения по сортировке остатков раскроя листового металла на деловые и неделовые. Многие CAD/CAM-системы и ERP-системы имеют общие интерфейсы взаимодействия для обмена информацией. Но интеграция этих программных средств, позволяющая решать более сложные задачи, связанные с формированием оперативных решений в условиях единичного и серийного производства, как правило, отсутствует. Поэтому задача исследования заключается в разработке методики сортировки остатков раскроя листового металла и построении программного комплекса на базе разработанной методики. Программный комплекс включает в себя систему поддержки принятия решений, помогающую в оперативном формировании оптимальных карт раскроя и оперативной и обоснованной сортировке остатков раскроя по группам деловых и неделовых остатков с целью увеличения коэффициента использования листового металла.

Основные функции программного комплекса в части управления материальными ресурсами листового металла

Программный комплекс, объединяя в себе CAD/CAM-систему, ERP-систему и СППР сортировки остатков раскроя, выполняет ряд функций. Типовой функционал CAD/CAM-системы позволяет:

- оптимальным образом размещать изделия задания на раскрой на листах метал-

ла, при этом учитывая технологические ограничения раскроя, например, расстояния до края листа и между деталями;

- осуществлять автоматическое программирование обработки листа за счет встроеного постпроцессора;

- сохранять в базу данных неиспользованную часть листа для проверки на возможность его дальнейшего использования;

- формировать выходные документы (карту раскроя листа, спецификацию раскроя листа, спецификацию задания на раскрой).

Подсистема «Управление заготовительным производством», входящая в состав ERP-системы, расширяет базовый функционал программного комплекса и позволяет:

- производить штрихкодирование изделий и деловых остатков;

- организовывать учет остатков листового металла после раскроя и материальных затрат на производство продукции;

- оценивать затраты на резку листового металла;

- автоматически подбирать цельные листы металла и деловые остатки, подходящие для задания на раскрой;

- проводить анализ процесса раскроя за произвольный период средствами специализированных отчетов.

Как СППР по сортировке остатков раскроя, подсистема «Управление заготовительным производством»:

- оказывает помощь пользователям в принятии решения по сортировке остатков на деловые и неделовые;

- производит переоценку деловых остатков на складе с целью выявления остатков, хранение которых становится экономически не выгодно.

Реализация программного комплекса

Методика сортировки остатков раскроя листового металла в группы деловых и неделовых остатков раскрывается далее в статье через описание алгоритмов функционирования программных модулей, составляющих программный комплекс и реализующих предложенную методику [7, 8].



Рис. 1. Структурная схема программного комплекса
Fig. 1. The structural diagram of the software package

С функциями, связанными с рациональным раскроем, справляется CAD/CAM-система, входящая в программный комплекс. Реализация оперативного, управленческого и бухгалтерского учета, реализация соответствующей отчетности и планирования легла на ERP-систему. Причем качество предоставляемой пользователю отчетности зависит от данных, собираемых ERP-системой, интерфейсов взаимодействия, входящих в состав комплекса систем, а также от алгоритмов работы СППР. Общая структурная схема программного комплекса представлена на рис. 1.

Графическое изображение взаимодействия информационных систем и пользователей в рамках процесса идентификации и сортировки остатков раскроя листового металла с помощью UML Activity diagram представлено на рис. 2 [8, 9].

Взаимодействие систем осуществляется с учетом экономических аспектов сортировки и хранения остатков раскроя. Итоговое отнесение остатков раскроя в группу деловых или неделовых материальных ресурсов осуществляется на основании сравнения изменения чистого дохода в зависимости от производства продукции из остатков раскроя или их реализации внешним организациям.

Далее более подробно рассмотрим входящие в программный комплекс системы. На рис. 3 представлена структурная схема CAD/CAM-системы в рамках реализованного программного комплекса.

Структура CAD/CAM-системы

Структуру типовой CAD/CAM-системы составляют отдельные программные модули. Модуль проектирования изделий позволяет создавать детали и заготовки на основе шаблонов или полностью ручным формированием контуров. База данных CAD/CAM-системы синхронизируется с базой данных разрабатываемой подсистемы «Управление заготовительным производством» ERP-системы. В базе данных CAD/CAM-системы хранятся изделия, которые могут быть импортированы из других систем, комплекты изделий, цельные листы исходного материала, возвратные остатки. Модуль раскроя листового металла позволяет осуществить рациональное размещение деталей и заготовок на цельном листе металла или листе возвратного остатка по заданию на раскрой. После того как изделия размещены на листе, модуль CAM программирует обработку. По завершении обработки CAD/CAM-система формирует выходные документы: карту раскроя, спецификацию раскроя и спецификацию задания на раскрой.

Структура ERP-системы

Структура ERP-системы значительно сложнее. Для более подробной отчетности и более точного прогнозирования в части использования деловых остатков, чем могут предложить методы, представленные в существующих ERP-системах, типовой функционал используемой ERP-системы был доработан.

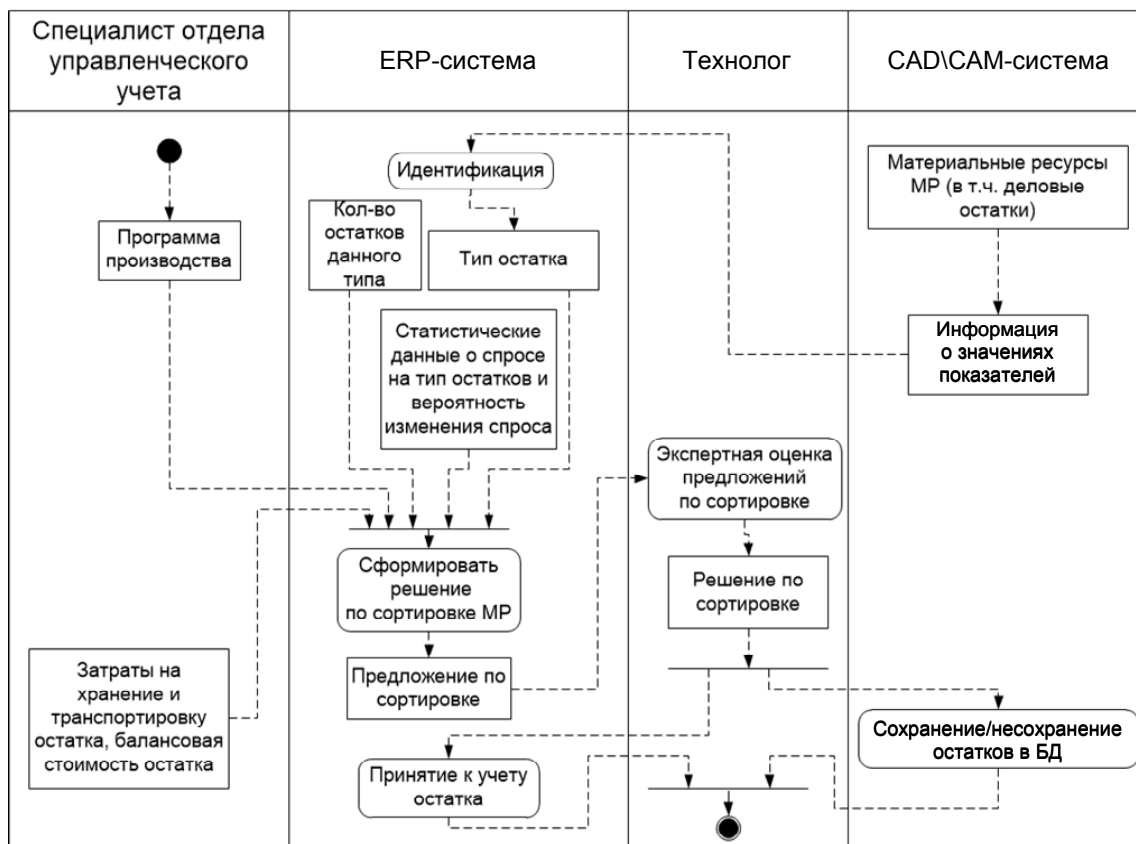


Рис. 2. Процесс идентификации и сортировки остатков раскроя листового металла
Fig. 2. The process of identification and sorting of sheet metal cutting residues

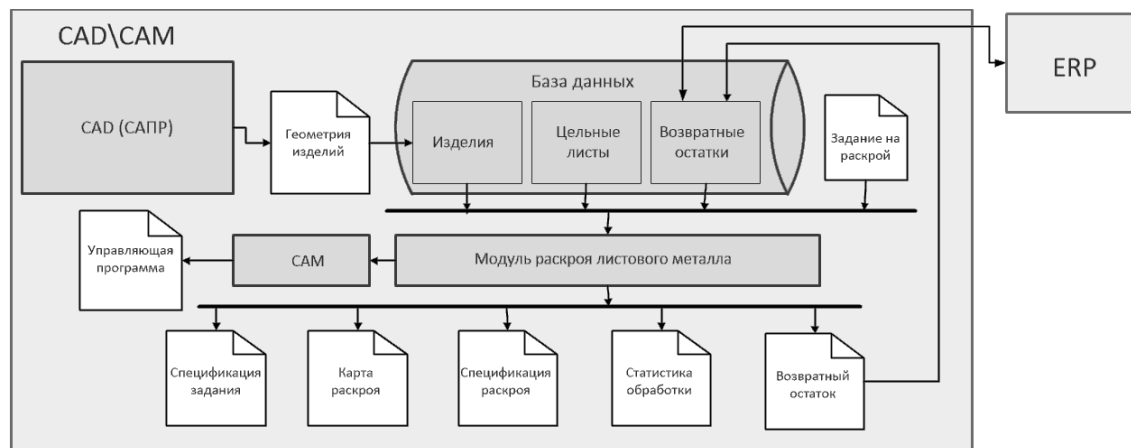


Рис. 3. Структурная схема CAD\CAM-системы в рамках разрабатываемого программного комплекса
Fig. 3. The structural diagram of the CAD\CAM-system in the framework of the developed software package

Многие ERP-системы имеют возможность конфигурирования, т. е. возможность разработки прикладных решений и модулей на базе платформы, на которой построена сама ERP-система. Состав прикладных механизмов платформы ERP-систем, как прави-

ло, ориентирован на решение задач автоматизации учета и управления предприятием и позволяет разработчику решать самый широкий круг задач складского, бухгалтерского, управленческого учета, анализа данных и управления на уровне бизнес-процессов.

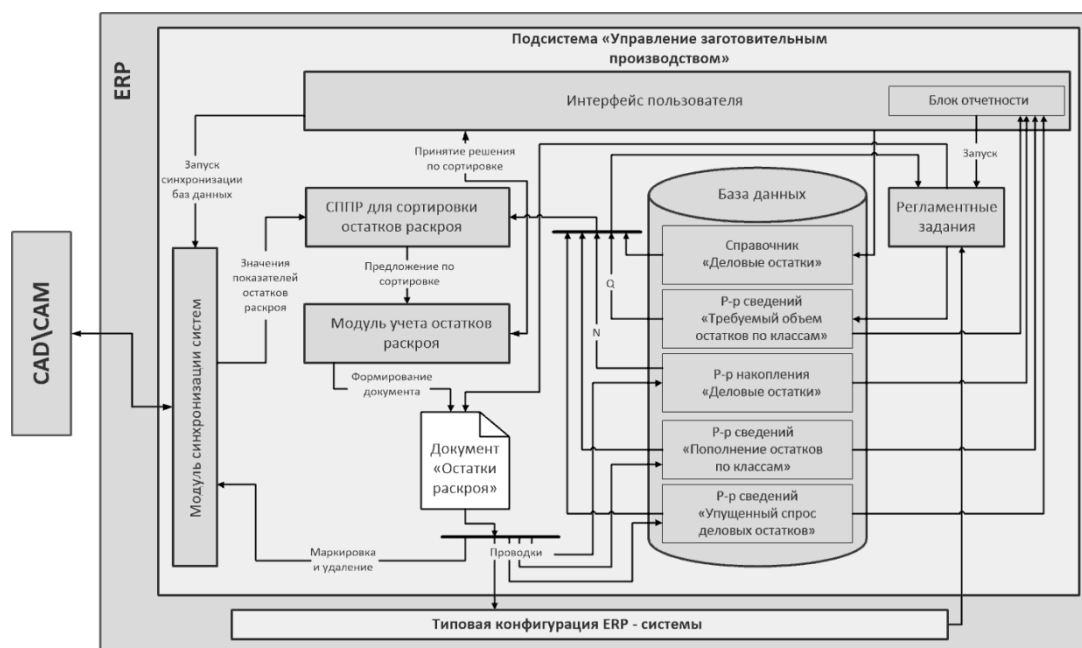


Рис. 4. Структурная схема ERP-системы в рамках разрабатываемого программного комплекса
Fig. 4. The structural diagram of the ERP-system in the framework of the developed software complex

ERP-система данного программного комплекса строится на базе технологической платформы «1С: Предприятие 8». Для реализации предложений, в целях повышения коэффициента использования листового металла, типовая структура ERP-системы дополнена подсистемой, реализующей управление заготовительным производством. Дополнительная подсистема «Управление заготовительным производством» позволяет расширить функциональные возможности программного комплекса в части реализации СППР сортировки остатков раскроя листового металла на деловые и неделовые, переоценки остатков, хранящихся на складах, а также взаимодействия CAD/CAM и ERP-систем. Структурно подсистема «Управление заготовительным производством» строится из взаимодействующих программных модулей, каждый из которых выполняет свою собственную функцию. Структура подсистемы «Управление заготовительным производством» и ее место в ERP-системе представлены на рис. 4.

Рассмотрим модули, из которых строится подсистема «Управление заготовительным производством».

Модуль синхронизации систем. Назначение модуля синхронизации систем сводится к обеспечению взаимодействия между CAD/CAM-системой и ERP-системой, путем реализации двусторонней синхронизации и обеспечения копирования актуальных изменений из одной системы в другую. Данный модуль позволяет обратиться к базе данных CAD/CAM-системы и выбрать из нее значения показателей тех возвратных остатков, по которым еще не принято решение: являются ли они деловыми или неделовыми. Значения показателей таких остатков модуль синхронизации передает модулю СППР для дальнейшей обработки и формирования рекомендации по сортировке. С помощью модуля синхронизации также осуществляется удаление из базы данных CAD/CAM-системы тех возвратных остатков, которые были признаны СППР неделовыми, и производится маркировка деловых остатков в CAD/CAM-системе в соответствии с их маркировкой в ERP-системе. С помощью данного модуля по запросу пользователя производится сравнение данных в информационных системах программного комплекса, что позволяет

выявить и своевременно обработать обнаруженные расхождения данных в CAD/CAM- и ERP-системах.

Распространенный и достаточно простой способ взаимодействия систем — использование общей базы данных. В условиях реализуемого программного комплекса наиболее подходящим является совместное использование базы данных CAD/CAM-системы самой CAD/CAM-системой и ERP-системой. Такой способ реализации взаимодействия характеризуется простотой реализации, не требующей включения в разрабатываемый комплекс дополнительных программных средств, надежностью и пропускной способностью канала связи информационных систем.

База данных. База данных подсистемы «Управление заготовительным производством» входит в состав базы данных ERP-системы. Рассмотрение базы данных можно провести, проанализировав только логические элементы базы, доступные пользователю, не вдаваясь в подробности относительно способа хранения данных, т. к. структуру данных на физическом уровне определяет сама платформа, на которой строится конфигурация ERP-системы. Множество таблиц и полей таблиц базы данных определяется типом объекта метаданных верхнего уровня (константы, справочники, документы, перечисления, регистры и др.), а также составом и типами подчиненных им объектов метаданных.

Для учета и регистрации хозяйственных операций на предприятиях применяют разнообразные по форме, содержанию и способу отображения информации учетные регистры. Обычно под регистром понимают различные виды таблиц, в которые записываются данные с первичных документов. В дальнейшем данные регистров используются для анализа хозяйственной деятельности предприятия. По данным регистров строятся отчеты, которые суммируют эти данные и показывают пользователю. В некоторых ERP-системах, в частности, построенных на платформе «1С: Предприя-

тие», документы записывают свои данные в такие регистры, которые сами суммируют результаты, чтобы отчет отобразил заранее посчитанные итоги. Существуют различные виды реализации регистров. Остановимся лишь на видах, реализующихся в базе данных подсистемы «Управление заготовительным производством». Регистры сведений позволяют хранить произвольные данные в разрезе нескольких измерений. Периодические регистры сведений позволяют хранить данные не только в разрезе указанных измерений, но и в разрезе времени. Регистры накопления реализуют учет движения средств, позволяют накапливать числовые данные в разрезе нескольких измерений. В разрабатываемом программном комплексе в регистре сведений накапливается информация об остатках товаров в разрезе номенклатуры и склада.

В рамках реализации подсистемы «Управление заготовительным производством» в состав конфигурации ERP-системы добавляется регистр накопления «Деловые остатки», который служит для определения количества деловых остатков определенного класса на определенный момент времени. Кроме того, по данному регистру можно узнать статистику потребления остатков определенного класса за выбранный период. Записи в данном регистре осуществляются в разрезе классов остатков раскроя, периода и вида движения по регистру: приход или расход.

Кроме регистра накопления в систему добавляются регистры сведений. В периодическом регистре сведений «Упущенный спрос деловых остатков» регистрируются те случаи, когда рациональнее использовать остаток определенного класса вместо цельного листа металла. Решение о рациональности использования остатка принимает CAM/CAD-система. Измерениями регистра являются период и класс остатка, а ресурсом — количество.

В периодическом регистре сведений «Пополнение классов остатков» фиксируется получившееся в результате раскроя количество остатков каждого класса.

Периодический регистр сведений «Требуемый объём остатков» необходим для определения требуемого в конкретный момент времени количества деловых остатков каждого класса. Состав измерений и ресурсов у регистров «Пополнение классов остатков» и «Требуемый объём остатков по классам» такой же, как у регистра «Упущенный спрос деловых остатков».

Помимо регистров в базу данных подсистемы «Управление заготовительным производством» входит справочник «Остатки раскроя», который служит для связи классов остатков с исходными листами металла и для хранения диапазона значений показателей, характеризующих конкретный класс остатков.

Модуль СППР подсистемы «Управление заготовительным производством». Предназначен для сортировки остатков раскроя в группы деловых и неделовых остатков. Данный модуль с помощью встроенного алгоритма формирует предложение по сортировке, с которым пользователь может согласиться или не согласиться. На вход СППР от модуля синхронизации систем поступают значения параметров возвратных остатков. На основании этих значений СППР осуществляет деление возвратных

остатков на определенные классы. Далее для определения того, является ли остаток деловым, анализируется класс остатков, к которому принадлежит данный остаток. СППР принимает решение о том, что остаток является деловым в том случае, если требуемое количество остатков данного класса больше, чем имеющийся на складе запас. Требуемый объём остатков конкретного класса и количество остатков на складах модуль СППР получает, обратившись к базе данных модуля «Управление заготовительным производством». В базе данных копится статистика пополняемости склада остатками по классам и статистика упущенного спроса на остатки.

Также база данных подсистемы «Управление заготовительным производством» содержит таблицу значений требуемого объёма остатков каждого класса на определенную дату (регистр сведений «Требуемый объём остатков по классам»). Модуль СППР соотносит значение этой таблицы со значением количества остатков на складах и однозначно определяет, является ли возвратный остаток в настоящий момент времени деловым. Блок-схема формирования предложения по сортировке остатков раскроя представлена на рис. 5.

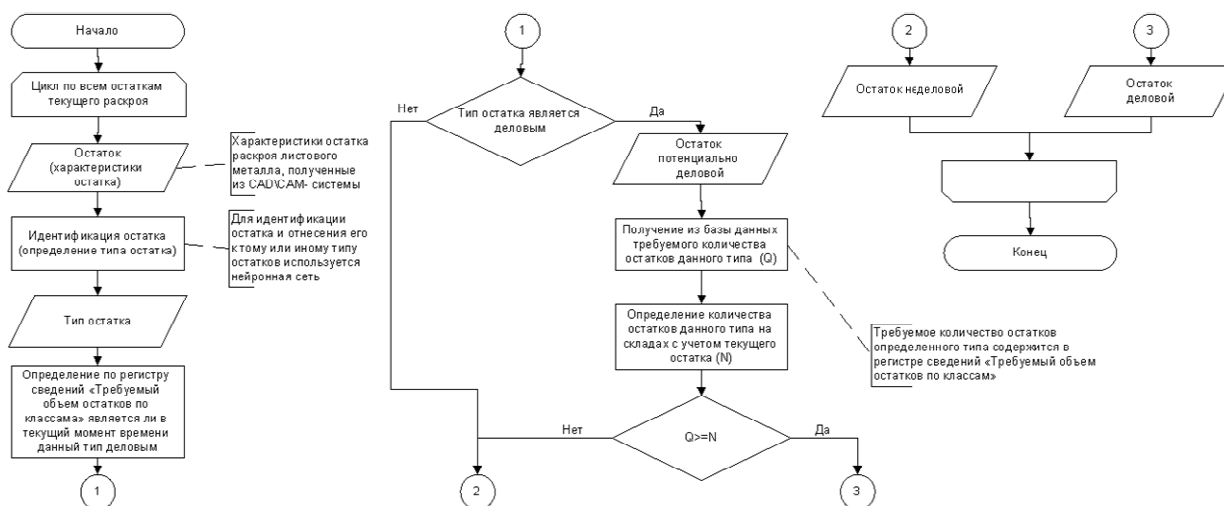


Рис. 5. Блок-схема формирования предложения по сортировке остатков раскроя модулем СППР

Fig. 5. The block diagram of the proposal for sorting residual cutting module DSS

После того как предложение по сортировке остатков сформировано СППР и одобрено пользователем, решение поступает в модуль учета остатков раскроя, где информация по остаткам обрабатывается дальше на основании принятого решения. Следует отметить, что при отнесении остатка в группы деловых и неделовых материальных ресурсов, итоговое решение необходимо принимать с учетом изменения чистого дохода предприятия.

Для отнесения остатков раскроя к тому или иному классу, СППР может использовать различные методы интеллектуального анализа данных: ассоциативные правила, деревья решений, генетические алгоритмы, системы обработки экспертных данных, искусственные нейронные сети и др. В реализуемом программном комплексе, в условиях нечеткости геометрических параметров делового остатка, подходящим методом классификации является использование нейронных сетей [9–12].

Модуль учета остатков раскроя. Служит для сохранения в системе остатков, признанных деловыми, и удаления неделовых остатков. Сначала одобренное пользователем решение по сортировке остатков поступает в данный модуль из модуля СППР. Результатом работы модуля учета остатков раскроя в ERP-системе является документ «Остатки раскроя», на основании которого в дальнейшем создаются типовые учетные документы для оприходования остатков раскроя. Документ «Остатки раскроя» входит в состав подсистемы «Управление заготовительным производством» и делает записи в регистр накопления «Деловые остатки», фиксирует в регистре сведений «Пополнение классов остатков», какое количество остатков каждого класса получилось в результате раскроя. После записи перечисленных регистров деловые остатки маркируются в базе данных CAD/CAM-системы в соответствии со своим классом, а неделовые остатки удаляются из CAD/CAM-системы. Маркировка и удаление в CAD/CAM-системе осуществляется при помощи модуля синхронизации систем.

На основании документа «Остатки раскроя» пользователь в дальнейшем создает учетные документы для принятия к учету деловых и неделовых остатков. Создание типовых учетных документов на основании документа «Остатки раскроя» избавляет пользователя от повторного ввода данных, которые уже хранятся в информационной базе. Кроме того, документ «Остатки раскроя» позволяет сделать подсистему управления заготовительным производством независимой от типовой конфигурации ERP-системы, что значительно облегчает интегрирование данной подсистемы в ERP-систему за счет минимальных изменений типовых механизмов и объектов конфигурации ERP-системы. Используемый принцип минимального вмешательства в типовую конфигурацию позволяет внедрять подсистему управления заготовительным производством в ERP-системах различных предприятий. Таким образом обеспечивается тиражирование данного программного решения.

Модуль регламентных заданий. Помимо перечисленных модулей, в состав подсистемы «Управление заготовительным производством» входит модуль регламентных заданий. Регламентные задания — это процедуры, которые выполняются по заданному расписанию, либо по запросу пользователя.

В подсистеме «Управление заготовительным производством» реализуется два регламентных задания. Первое регламентное задание — это «Расчет потребности в остатках каждого класса». Его целесообразно выполнять при планировании производства. Данное задание делает запись в регистр сведений «Требуемый объем остатков по классам». Второе регламентное задание — «Переоценка хранящихся деловых остатков». Переоценка хранящихся остатков необходима, поскольку может изменяться номенклатура производимых заготовок, а также затраты на хранение остатка. Данное регламентное задание формирует отчет об экономической целесообразности хранения остатков различных классов.

Интерфейс пользователя. Возможность централизованного администрирования подсистемы «Управление заготовительным производством» предоставляет интерфейс пользователя. Он позволяет пользователю запускать регламентные задания, настраивать расписание их выполнения, реализует диалоговый режим работы модуля СППР и формирование документа «Остатки товаров». В интерфейс пользователя также входит блок отчетности, позволяющий производить анализ операций заготовительного производства на основании записей регистров.

Заключение

В результате исследования рассмотрены способы увеличения коэффициента использования листового металла.

Предложена методика сортировки остатков раскроя листового металла в группы деловых и неделовых материальных ресурсов, которая предполагает идентификацию остатков на основе оценки значений показателей, характеризующих остатки раскроя.

На основе предложенной методики спроектирован программный комплекс

управления материальными ресурсами листового металла с элементами СППР, помогающий снизить удельные технологические затраты и оптимизировать структуру затрат машиностроительного предприятия. Входящая в состав программного комплекса СППР позволяет принимать оперативные и обоснованные управленческие решения при отнесении остатков раскроя листового металла в группу деловых или неделовых остатков. Апробация методики сортировки материальных ресурсов после раскроя показала увеличение коэффициента использования листового металла в среднем с 0,8 до 0,9. Следовательно, при производстве одинакового объема продукции исходных материальных ресурсов листового металла требуется меньше примерно на 11 %, а удельные операционные затраты на листовую металл снижаются примерно на 10 % (при учете неделовых материальных ресурсов по цене металлолома). Подобные значения показателей эффективности ожидаются и при промышленной эксплуатации программного комплекса в условиях машиностроительных предприятий по производству емкостного оборудования.

СПИСОК ЛИТЕРАТУРЫ

1. Канторович Л.В., Залгаллер В.А. Рациональный раскрой промышленных материалов. Новосибирск: Наука, 1971. 300 с.
2. Мухачева Э.А., Верхотуров М.А., Мартынов В.В. Модели и методы расчета раскроя упаковки геометрических объектов. Уфа: УГАТУ, 1998. 216 с.
3. Мухачева Э.А. Рациональный раскрой промышленных материалов. Применение АСУ. М.: Машиностроение, 1984. 176 с.
4. Kartak V.M., Mesyagutov M.A., Mukhacheva E.A., Filippova A.S. Local search of orthogonal packings using the lower bounds // Automation and Remote Control. 2009. Vol. 70(6). Pp. 1054–1066.
5. Валиахметова Ю.И., Филиппова А.С. Теория оптимального использования ресурсов Л.В. Канторовича в задачах раскроя-упаковки: обзор и история развития методов решения // Вестник Уфимского государственного авиационного технического университета. 2014. № 1. С. 186–197.
6. Петунин А.А. Автоматический выбор метода расчета фигурного раскроя с использованием сравнительного анализа алгоритмов // Известия Томского политехнического университета. Инжиниринг георесурсов. 2010. № 5 (316). С. 169–171.
7. Смирнов А.А. Экономические аспекты принятия управленческих решений при идентификации материальных ресурсов листового металла на машиностроительном предприятии // Научно-технические ведомости СПбГПУ. Экономические науки. 2017. № 6. С. 195–204. DOI 10.18721/JE.10618
8. Смирнов А.А., Шичков А.Н. Организация раскроя листового металла с учетом деловых материальных ресурсов на машиностроительном предприятии // Вестник ЮРГТУ (НПИ). 2017. № 4. С. 22–35. DOI: 10.17213/2075-2067-2017-4-22-35
9. Швецов А.Н., Суконщиков А.А., Кочкин Д.В., Андрианов И.А. Ситуационные интеллектуальные системы поддержки принятия решений. Курск: Университетская книга, 2018. 250 с.
10. Дохтаева И.А., Суконщиков А.А. Современные методы интеллектуального анализа данных в СППР // Информатизация инженерного

образования: Труды Междунар. науч.-практ. конф. М.: Изд. дом МЭИ, 2016. С. 312–315.

11. Курейчик В.М., Сафроненкова И.Б. Разработка архитектуры СППР по выбору методов решения задач компоновки // Информационные технологии. 2017. № 10(23). С. 42–48.

12. Задорожный А.М., Чуваков А.В. Положительные аспекты использования СППР, принципы работы и классификация // Теория и практика современной науки. 2017. № 6(24). С. 306–309.

13. Skorodumov P.V. Modelling of economic systems with Petri nets // Economic and Social

Changes: Facts, Trends, Forecast. 2014. No. 4. Pp. 253–259.

14. Wascher G., Haubner H., Schumann H. An improved typology of cutting and packing problems // European Journal of Operational Research. 2007. Vol. 183(3). Pp. 1109–1130.

15. Улитин А.В., Суконщиков А.А. Применение аппарата нейро-нечетких сетей Петри для моделирования СППР // Сб. науч. трудов XII Междунар. науч. конф. Современные инструментальные системы, информационные технологии и инновации. Курск: Университетская книга, 2015. С. 160–163.

Статья поступила в редакцию 27.01.2019.

REFERENCES

1. Kantorovich L.V., Zalgaller V.A. *Ratsionalnyy raskroy promyshlennyykh materialov* [Rational cutting of industrial materials]. Novosibirsk: Nauka Publ., 1971, 300 p. (rus)

2. Mukhacheva E.A., Verkhoturov M.A., Martynov V.V. *Modeli i metody rascheta raskroya upakovki geometricheskikh obyektov* [Models and calculation methods for cutting packaging of geometric objects]. Ufa: UGATU Publ., 1998, 216 p. (rus)

3. Mukhacheva E.A. *Ratsionalnyy raskroy promyshlennyykh materialov. Primeneniye ASU* [Rational cutting of industrial materials. ACS application]. Moscow: Mashinostroyeniye Publ., 1984, 176 p. (rus)

4. Kartak V.M., Mesyagutov M.A., Mukhacheva E.A., Filippova A.S. Local search of orthogonal packings using the lower bounds. *Automation and Remote Control*, 2009, Vol. 70(6), Pp. 1054–1066.

5. Valiakhmetova Yu.I., Filippova A.S. Teoriya optimalnogo ispolzovaniya resursov L.V. Kantorovicha v zadachakh raskroya-upakovki: obzor i istoriya razvitiya metodov resheniya [Theory of optimum resource utilization by L.V. Kantorovich in cutting-packing problems: overview and history of development of solving methods]. *Vestnik Ufimskogo gosudarstvennogo aviatsionnogo tekhnicheskogo universiteta*, 2014, No. 1, Pp. 186–197. (rus)

6. Petunin A.A. Avtomaticheskii vybor metoda rascheta figurnogo raskroya s ispolzovaniyem sravnitel'nogo analiza algoritmov [Automatic selection of a method for calculating shape cutting using a comparative analysis of algorithms]. *Izvestiya Tomskogo politekhnicheskogo universiteta. Inzhiniring georesursov*, 2010, No. 5 (316), Pp. 169–171. (rus)

7. Smirnov A.A. Ekonomicheskiye aspekty prinyatiya upravlencheskikh resheniy pri identifikatsii materialnykh resursov listovogo metalla na

mashinostroitel'nom predpriyatii [Economic aspects of making managerial decisions in the identification of material resources at an engineering enterprise]. *Nauchno-tekhnicheskiye vedomosti SPbGPU. Ekonomicheskiye nauki* [St. Petersburg State Polytechnical University Journal. Economics], 2017, No. 6, Pp. 195–204. (rus) DOI: 10.18721/JE.10618

8. Smirnov A.A., Shichkov A.N. Organizatsiya raskroya listovogo metalla s uchetom delovykh materialnykh resursov na mashinostroitel'nom predpriyatii [Organization of sheet metal cutting taking into account business material resources at a machine-building enterprise]. *Vestnik YuRGU (NPI)*, 2017, No. 4, Pp. 22–35. (rus) DOI: 10.17213/2075-2067-2017-4-22-35

9. Shvetsov A.N., Sukonshchikov A.A., Kochkin D.V., Andrianov I.A. *Situatsionnyye intellektualnyye sistemy podderzhki prinyatiya resheniy* [Situational Intelligent Decision Support Systems]. Kursk: Universitetskaya kniga Publ., 2018, 250 p. (rus)

10. Dokhtayeva I.A., Sukonshchikov A.A. Sovremennyye metody intellektual'nogo analiza dannykh v SPPR [Modern methods of data mining in DSS]. *Informatizatsiya inzhener'nogo obrazovaniya: trudy Mezhdunar. nauch.-prakt. konf.* [Proceedings of the Intern. Scientific and Practical Conference on Informatization of Engineering Education]. Moscow: MEI Publ., 2016, Pp. 312–315. (rus)

11. Kureychik V.M., Safronenkova I.B. Razrabotka arkhitektury SPPR po vyboru metodov resheniya zadach komponovki [Decision support system development for the choice of problem-solving procedure of commutation circuit partitioning]. *Informatsionnyye Tekhnologii* [Information Technology], 2017, No. 10(23), Pp. 42–48. (rus)

12. Zadorozhnyy A.M., Chuvakov A.V. Polozhitel'nyye aspekty ispolzovaniya SPPR,

printsipy raboty i klassifikatsiya [Positive aspects of DSS usage, method of procedure and classification of DSS]. *Teoriya i Praktika Sovremennoy Nauki* [Theory and Practice of Modern Science], 2017, No. 6(24), Pp. 306–309. (rus)

13. **Skorodumov P.V.** Modelling of economic systems with Petri nets. *Economic and Social Changes: Facts, Trends, Forecast*, 2014, No. 4, Pp. 253–259.

14. **Wascher G., Haubner H., Schumann H.** An improved typology of cutting and packing problems.

European Journal of Operational Research, 2007, Vol. 183(3), Pp. 1109–1130.

15. **Ulitin A.V., Sukonshchikov A.A.** Primeneniye apparata neyro-nechetkikh setey Petri dlya modelirovaniya SPPR [Application of the apparatus of neuro-fuzzy Petri nets for modeling DSS]. *Proceedings of the 12th International Scientific Conf. on Modern Instrumental Systems, Information Technology and Innovation*. Kursk: Universitetskaya kniga Publ., 2015, Pp. 160–163. (rus)

Received 27.01.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

УЛИТИН Анатолий Вячеславович

ULITIN Anatoly V.

E-mail: ulitko_onotole@mail.ru

СМИРНОВ Артём Алексеевич

SMIRNOV Artem A.

E-mail: smirnov.artem.a@yandex.ru

DOI: 10.18721/JCSTCS.12405
УДК 621.391.8

ИТЕРАТИВНЫЙ ПРИЁМ МНОГОЧАСТОТНЫХ СИГНАЛОВ С НЕОРТОГОНАЛЬНЫМ ЧАСТОТНЫМ УПЛОТНЕНИЕМ В КАНАЛАХ С ЗАМИРАНИЯМИ

А.В. Рашич, С.В. Горбунов

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Для приёма кодированных SEFDM-сигналов в каналах с частотно-селективными замираниями предложен итеративный алгоритм приёма, основанный на совместном применении ZF-эквалайзера эквивалентного OFDM-сигнала, алгоритма BCJR для демодуляции и декодирования сверточного кода. Проведен анализ помехоустойчивости приёма SEFDM-сигнала в многолучевых каналах с замираниями EPA 5Hz, EVA 70Hz, ETU 70Hz. Оценка помехоустойчивости приёма SEFDM-сигналов в условиях канала с замираниями выполнялась методами имитационного моделирования при условии идеального знания канала на стороне приёмного устройства. Описанный метод приёма обеспечивает низкую вероятность ошибки при относительно небольшом энергетическом проигрыше по сравнению с приёмом аналогичных ортогональных систем сигналов. Вычислительная сложность предложенного метода растет линейно при увеличении частоты. Разработанный метод применим как для малого, так и для большого числа поднесущих в сигнале.

Ключевые слова: OFDM, NOFDM, SEFDM, неортогональные многочастотные сигналы, 5G, BCJR, каналы с замираниями.

Ссылка при цитировании: Рашич А.В., Горбунов С.В. Итеративный приём многочастотных сигналов с неортогональным частотным уплотнением в каналах с замираниями // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 58–71. DOI: 10.18721/JCSTCS.12405

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

ITERATIVE RECEIVER OF SEFDM SIGNALS IN FADING CHANNELS

A.V. Rashich, S.V. Gorbunov

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

An iterative receiver of SEFDM signal for the case of fading channel is proposed. The algorithm is based on the BCJR algorithm demodulator and soft decision decoder of convolutional code. Also the equivalent representation of SEFDM as OFDM with

reduced number of subcarriers is used for equalization. The proposed algorithm provides small BER with relatively small energy loss compared to equivalent OFDM system. A study of SEFDM signals behavior in frequency selective fading channel is done. The proposed algorithm have computational complexity linearly increasing with the number of subcarriers and are applicable for any number of subcarriers in the multicarrier signal.

Keywords: OFDM, NOFDM, SEFDM, Multicarrier FTN, 5G, BCJR, fading channels.

Citation: Rashich A.V., Gorbunov S.V. Iterative receiver of SEFDM signals in fading channels. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 58–71. DOI: 10.18721/JCSTCS.12405

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Большое распространение в современных широкополосных беспроводных системах передачи данных (WiFi-a, g, n, ac, ax, LTE, DVB-T2, 5G) получили многочастотные сигналы с ортогональным частотным уплотнением (Orthogonal Frequency Division Multiplexing — OFDM). Для дальнейшего увеличения скорости передачи необходимо повышать частотную эффективность систем сигналов. В качестве одной из альтернатив OFDM для использования в стандартах связи следующих поколений рассматривается метод неортогонального частотного уплотнения, получивший в англоязычной научной литературе название Spectrally Efficient Frequency Division Multiplexing (SEFDM). SEFDM-сигналы формируются из ортогональных сигналов путем уменьшения расстояния между поднесущими частотами в $1/\alpha$ раз, где α — коэффициент частотного уплотнения поднесущих частот, тем самым повышая спектральную эффективность [1, 10].

В отличие от OFDM, в случае SEFDM-сигналов возникает взаимная интерференция между поднесущими (внутрисимвольная интерференция), вследствие которой использование алгоритмов, разработанных для приёма OFDM-сигналов (когда демодуляция выполняется независимо на каждой поднесущей), приводит к наличию несократимой средней вероятности ошибки на бит (СВО на бит) $< 10^{-2}$ для

SEFDM-сигналов с параметром $\alpha < 7/8$. Следовательно, внутрисимвольная интерференция на приёмной стороне должна быть устранена или снижена до приемлемого значения. Известными подходами для решения задачи компенсации внутрисимвольной интерференции SEFDM-сигналов в канале с аддитивным белым гауссовским шумом (АБГШ) являются алгоритм последовательного подавления помех [2], алгоритм на основе принципов сферического декодирования [3], линейные алгоритмы, заключающиеся в инверсии матрицы, описывающей внутрисимвольную интерференцию сигнала [4].

В работе [5] предложен демодулятор, основанный на алгоритме BCJR (назван по имени авторов Bahl, Cocke, Jelinek и Raviv) [6], особенностью которого является то, что в процессе приёма SEFDM-сигнала учитывается интерференция от ограниченного количества соседних поднесущих. Алгоритм сферического декодирования [11, 12], изначально предложенный для декодирования помехоустойчивых кодов, решает систему линейных уравнений методом максимального правдоподобия при заданных ограничениях на область поиска [13, 14]. Использование такого алгоритма для демодуляции SEFDM-сигналов обеспечивает результаты, близкие к демодуляции по методу максимального правдоподобия, но имеет высокую сложность и, таким обра-

зом, применимо только для SEFDM-сигналов с небольшим (5–10) количеством поднесущих. Алгоритмы последовательного подавления помех и линейные алгоритмы имеют низкую вычислительную сложность, но при этом их использование ведет к наличию несократимой средней вероятности ошибки на бит (СВО на бит) $\approx 10^{-3}$ для сигналов с параметром $\alpha < 4/5$. Демодулятор, основанный на алгоритме BCJR, обеспечивает приемлемую производительность и низкую СВО на бит, обладая вычислительной сложностью, растущей линейно от количества поднесущих.

Приём SEFDM-сигналов в условиях канала с замираниями рассмотрен в [7–9]. Работы [7, 8] обобщают использование алгоритма сферического декодирования для совместной эквализации и демодуляции SEFDM-сигнала. Более того, в указанных работах анализ был сделан только для упрощенных моделей канала с постоянными коэффициентами передачи импульсного отклика канала. В [9] предложен алгоритм эквализации, основанный на рассмотрении SEFDM-сигнала как OFDM-сигнала с меньшим в $1/\alpha$ количеством поднесущих (эквивалентный OFDM-сигнал) и новыми точками в сигнальном созвездии. Для обработки такого сигнала можно применять алгоритмы эквализации, предложенные для классических OFDM-сигналов. При этом для демодуляции сигнала использован алгоритм BCJR. Недостатком подхода, предложенного в [9], является наличие несократимой СВО на бит $\approx 10^{-3}$ для SEFDM-сигналов с параметром $\alpha < 5/8$.

В современных сетях связи, как правило, используются методы помехоустойчивого кодирования, позволяющие применять итеративные схемы приёма сигналов. В [5] предложена итеративная схема приёма SEFDM-сигнала, кодированного сверточным кодом, в канале с АБГШ, а итеративный приёмник основан на использовании алгоритма BCJR для демодуляции и декодирования и позволяет уменьшить энергетический проигрыш по сравнению с приёмом OFDM-сигнала с

тем же помехоустойчивым кодом до значений менее 1 дБ для СВО на бит $= 10^{-4}$.

При этом на данный момент не проанализировано сохраняет ли итеративная схема приёма свою эффективность в случае распространения сигнала в каналах с частотно-селективными замираниями. Более того, отсутствует анализ эффективности работы итеративного приёмника SEFDM-сигналов в сочетании с эквалайзером.

В настоящей статье выполнен анализ помехоустойчивости приёма SEFDM-сигналов, кодированного сверточным кодом, в канале с частотно-селективными замираниями, при использовании итеративного алгоритма приёма, основанного на совместном применении ZF-эквалайзера эквивалентного OFDM-сигнала, алгоритма BCJR для демодуляции и декодирования сверточного кода. При этом для моделирования итеративного приёма использованы каналы с замираниями, предложенные в стандарте связи LTE (Long Term Evolution).

Системная модель

Комплексную огибающую многочастотного сигнала с N поднесущими частотами на интервале времени $[0, T]$ в основной полосе частот можно записать следующим образом:

$$s(t) = \sum_{k=-N/2}^{N/2-1} C_k e^{j2\pi\Delta f k t}, t \in [0; T], \quad (1)$$

где C_k — манипуляционный символ k -й поднесущей, Δf — частотный разнос между соседними поднесущими.

В случае классического OFDM-сигнала частотный разнос $\Delta f = 1/T$; если $\Delta f = \alpha/T < 1/T$, где $0 < \alpha < 1$, то $s(t)$ является SEFDM-сигналом, а коэффициент α называется коэффициентом частотного уплотнения. Таким образом, для OFDM-сигналов $\alpha = 1$.

В дискретном времени комплексную огибающую многочастотного сигнала (1) можно представить в следующем виде:

$$s_n = \sum_{k=-N/2}^{N/2-1} C_k e^{j2\pi\frac{kn}{N}}, n = 0 \dots L-1, \quad (2)$$

где частота дискретизации $F_s = 1/\Delta t = N\Delta f$ и $T/\Delta t = N\alpha = L$, то есть количество отсчетов на длительности одного символа у дискретных SEFDM-сигналов в $1/\alpha$ раз меньше, чем у OFDM-сигналов.

При этом для уменьшения влияния наложения спектров соседних по частоте сигналов используются защитные интервалы по частоте:

$$C_k = 0 \text{ при } k \in [0; N_{GL_left} - 1] \text{ и } k \in [N - N_{GL_right}; N - 1], \quad (3)$$

где N_{GL_left} — количество неиспользуемых поднесущих «слева» в области отрицательных частот, N_{GL_right} — количество неиспользуемых поднесущих «справа» в области положительных частот.

Из (2) следует, что отсчеты многочастотного сигнала s_n могут быть получены в результате выполнения обратного дискретного преобразования Фурье (ОДПФ) размерности N от последовательности модуляционных символов $\{C_k\}_{k=-N/2}^{N/2-1}$ с последующим усечением получаемой на выходе ОДПФ последовательности до L отсчетов (рис. 1) [15].

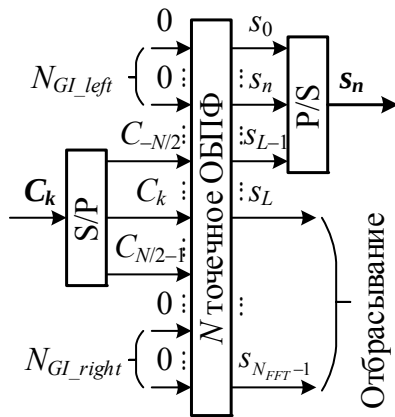


Рис. 1. Структурная схема алгоритма формирования SEFDM-сигналов

Fig. 1. The structure of SEFDM signal generator

Выражение (2) можно представить в виде умножения вектор-столбца $\vec{c} = \{C_k\}_{k=-N/2}^{N/2-1}$ на матрицу $F_{L \times N}^{-1}$, где $F_{L \times N}^{-1}$ — это матрица преобразования Фурье размерности N , выход которого усечен до длины L :

$$\vec{s} = F_{L \times N}^{-1} \times \vec{c}. \quad (4)$$

В случае многолучевого распространения сигнал достигает приёмника разными путями, каждый из которых является копией оригинала. Копии переданного сигнала на входе приёмника имеют разные фазовые сдвиги и среднюю мощность. В данной работе модель канала представлена линейным дискретным фильтром с конечной импульсной характеристикой h (отклик канала), на выходе которого к сигналу добавляется АБГШ. Таким образом, принятый SEFDM-сигнал может быть представлен как

$$r_i = \sum_{j=0}^{J-1} s_{i-j} h_j + w_i, \quad (5)$$

где J — количество путей распространения сигнала; h_j — коэффициенты импульсной характеристики канала; w_i — отсчеты АБГШ. В случае рэлеевских замираний h_j представляют собой комплексные случайные числа с гауссовским распределением.

Как и в случае с OFDM-сигналами, для борьбы с эффектами замираний сигнала в канале связи в системе с SEFDM-сигналами целесообразно использование защитного интервала в виде циклического префикса, который представляет собой копию последних L_{CP} отсчетов, добавленных перед сигналом. Его длина L_{CP} должна быть, по крайней мере, не меньше максимальной длительности импульсного отклика канала. Тогда выражение (5) может быть рассмотрено как циклическая свертка и записано следующим образом:

$$\vec{r} = \mathbf{H}_{L \times L} \times \vec{s} + \vec{w} = \mathbf{H}_{L \times L} \times F_{L \times N}^{-1} \times \vec{c} + \vec{w}, \quad (6)$$

где $\mathbf{H}_{L \times L}$ — это матрица размера $L \times L$, состоящая из циклически сдвинутых последовательностей $\vec{h} = \{h_j\}_{j=0}^{L-1}$. Системная модель представлена на рис. 2. Она состоит из блока помехоустойчивого кодирования, модулятора, формирователя многочастотного сигнала, многолучевого канала с АБГШ и приёмной части. В свою очередь приёмная часть состоит из двух основных блоков: блока компенсации эффектов канала (эквалайзер) (рис. 3) и блока демодуляции и декодирования (рис. 4).

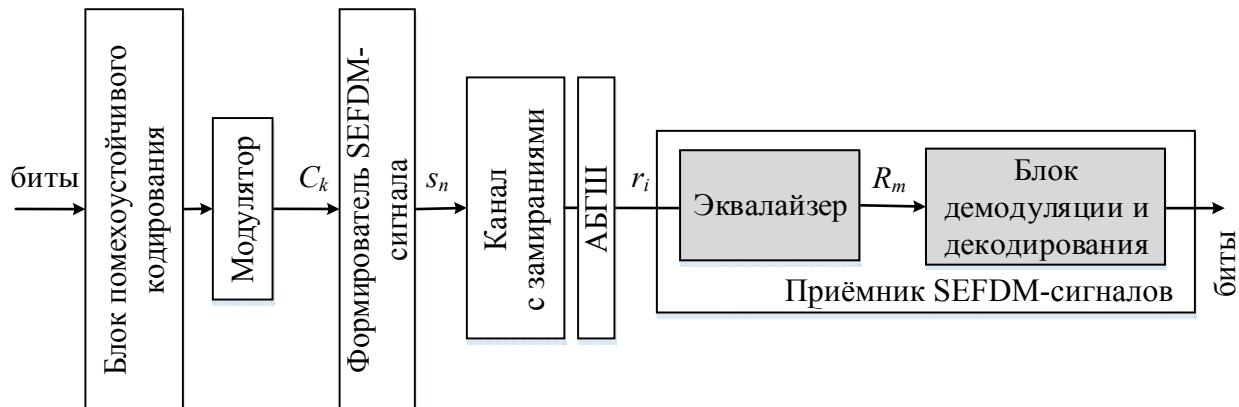


Рис. 2. Системная модель

Fig. 2. System model

Приёмник SEFDM-сигналов

Приём многочастотных сигналов осуществляется в частотной области по спектральным отсчетам $\{R_m\}_{m=-N/2}^{N/2-1}$, представляющим собой выход дискретного преобразования Фурье размерности N сигнала $\{s_n\}_{n=0}^{L-1}$ (обратная к (4) операция):

$$\vec{r}^{DFT} = \mathbf{F}_{N \times L} \times \mathbf{H}_{L \times L} \times \mathbf{F}_{L \times N}^{-1} \times \vec{c} + \vec{w}^{DFT}. \quad (7)$$

Выражение (7) в отсутствие канала с замираниями и шума выглядит как

$$\vec{r}^{DFT} = \mathbf{F}_{N \times L} \times \mathbf{F}_{L \times N}^{-1} \times \vec{c}. \quad (8)$$

То есть отсчеты $\vec{r}^{DFT} = \{R_m\}_{m=-N/2}^{N/2-1}$ могут быть представлены как

$$R_m = \sum_{i=0}^{L-1} s_n e^{j2\pi \frac{nm}{N}} = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} C_k \sum_{n=0}^{L-1} e^{j2\pi \frac{n(k-m)}{N}} = \sum_{k=-N/2}^{N/2-1} C_k g_{k-m}, \quad m = -N/2 \dots N/2-1, \quad (9)$$

где отсчеты $g_k = \sum_{n=0}^{L-1} e^{j2\pi \frac{nk}{N}}$ (ДПФ прямоугольного окна).

Таким образом, $\vec{r}^{DFT} = \{R_m\}_{m=-N/2}^{N/2-1}$ представляют собой свертку последовательности модуляционных символов $\vec{c} = \{C_k\}_{k=-N/2}^{N/2-1}$ и последовательности $\vec{g} = \{g_k\}_{k=-N/2}^{N/2-1}$. В [5] показано, что последовательность $\vec{g} = \{g_k\}_{k=-N/2}^{N/2-1}$, имеющая длину N , может быть аппроксимирована последовательностью меньшей дли-

ны K (типичные значения $K = 3, 5, 7$), а для демодуляции использован алгоритм, основанный на алгоритме BCJR.

В случае каналов с замираниями для использования алгоритмов демодуляции SEFDM-сигналов, разработанных для условий каналов с АБГШ, необходимо предварительно компенсировать эффект распространения в канале: свести матрицу $\mathbf{F}_{N \times L} \times \mathbf{H}_{L \times L} \times \mathbf{F}_{L \times N}^{-1}$ к $\mathbf{F}_{N \times L} \times \mathbf{F}_{L \times N}^{-1}$. Заметим, что SEFDM-сигнал ($\alpha < 1$) может быть представлен через эквивалентный OFDM-сигнал с L ортогональными поднесущими. При этом поднесущие модулированы комплексными символами $\{S_k\}_{k=0}^{L-1}$, представляющими собой L -точечное ДПФ сигнала $\{s_n\}_{n=0}^{L-1}$ [9, 16]:

$$S_k = \frac{1}{L} \sum_{n=0}^{L-1} s_n e^{j2\pi \frac{kn}{L}}, \quad k = 0 \dots (L-1); \quad (10)$$

$$s_n = \sum_{i=0}^{L-1} S_k e^{j2\pi \frac{nk}{L}}, \quad n = 0 \dots (L-1). \quad (11)$$

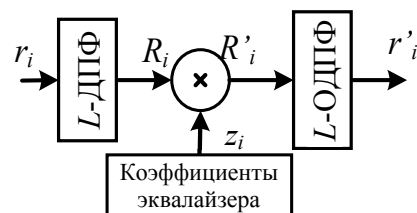


Рис. 3. Блок компенсации эффектов канала (эквалайзер)

Fig. 3. Equalizer

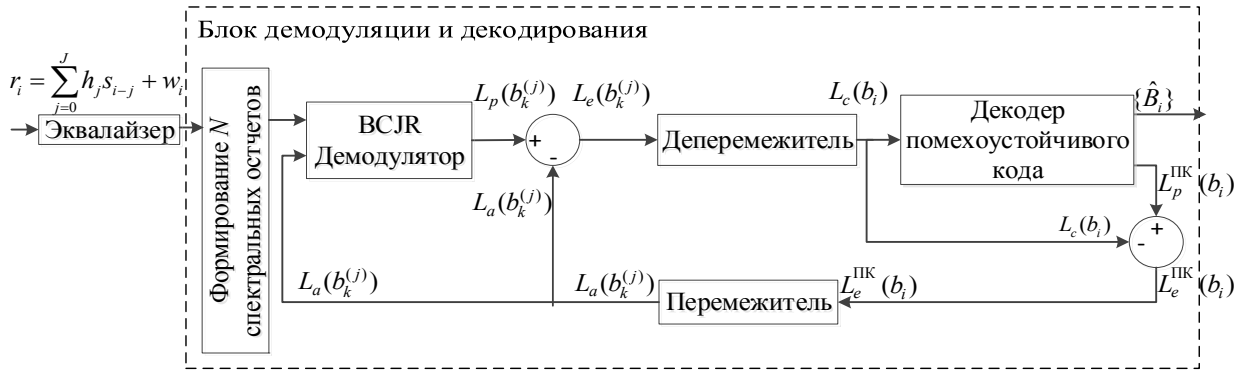


Рис. 4. Структурная схема разработанного итеративного алгоритма приёма SEFD-сигналов с использованием эквалайзера

Fig. 4. The proposed iterative receiver of SEFD-signals with equalizer

Используя представление SEFD-сигнала в виде эквивалентного OFDM-сигнала, в [9] предложен алгоритм компенсации эффектов канала, аналогичный используемым в системах связи с OFDM и осуществляющий независимую эквализацию каждой поднесущей эквивалентного OFDM-сигнала S_k . Таким образом, блок компенсации эффектов канала включает в себя обратное дискретное преобразование Фурье (переход к спектральным отсчетам $\{S_k\}_{k=0}^{L-1}$) и прямое дискретное преобразование Фурье размерности L (рис. 3) и состоит из следующих шагов:

ДПФ размерности L ;

эквализация (выравнивание частотной характеристики);

ОДПФ размерности L .

В данной статье коэффициенты эквалайзера рассчитываются при условии идеального знания канала на приёмном устройстве следующим образом:

$$Z_i = \frac{1}{H_i} = \frac{H_i^*}{|H_i|}, \quad (12)$$

где H_i представляет собой коэффициент передачи канала на i -й поднесущей S_n , который может быть найден как $H_i = \frac{1}{L} \sum_{n=0}^{L-1} h_n e^{j2\pi \frac{in}{L}}$, $i = 0 \dots (L-1)$.

Входом алгоритма BCJR является последовательность эквализированных спектральных отсчетов SEFD-символа [16, 17], а результатом работы алгоритма — последовательность мягких решений канальных битов в виде логарифмов отношения правдоподобия (Log-Likelihood Ratio — LLR [18]) $L_p(b_k^{(j)})$, где $b_k^{(j)}$ — j -й бит переданного манипуляционного символа C_k . Далее $L_p(b_k^{(j)})$ используются в качестве мягких решений на входе декодера.

Вероятность ошибки на выходе декодера можно уменьшить с помощью перехода к итеративной схеме приёма [19, 20]. В такой схеме на каждой следующей итерации решение с выхода декодера поступает на вход блока демодуляции, где выступает в роли априорной информации о канальных битах. Разработанный демодулятор на основе алгоритма BCJR принимает информацию об априорных вероятностях канальных битов $L_a(b_k^{(j)})$.

При отсутствии априорной информации (перед первой итерацией итеративного алгоритма приёма) данные отношения определяются нулевым значением.

Одной из основных проблем алгоритмов итеративного приёма является проблема коррелированности мягких решений канальных битов, повышающаяся при

увеличении количества итераций. Для снижения коррелированности мягких решений в итеративном алгоритме приёма SEFDM-сигналов используются перемежения/деперемежения и внешняя информация о канальных битах в виде логарифмов отношения правдоподобия канальных битов на входе декодера вместо $L_p(b_k^{(j)})$, представляющие собой разность между $L_p(b_k^{(j)})$ и $L_a(b_k^{(j)})$:

$$L_e(b_k^{(j)}) = L_p(b_k^{(j)}) - L_a(b_k^{(j)}). \quad (13)$$

При отсутствии априорной информации на первой итерации $L_a(b_k^{(j)}) = 0$ и $L_e(b_k^{(j)}) = L_p(b_k^{(j)})$. Для алгоритмов декодирования помехоустойчивых кодов внешние логарифмы отношения правдоподобия вычисляются в соответствии с выражением:

$$L_e^{\text{ПК}}(b_i) = L_p^{\text{ПК}}(b_i) - L_c(b_i). \quad (14)$$

То есть внешняя информация представляет собой информацию, полученную в процессе демодуляции SEFDM-сигналов или декодирования без учета априорной информации на входе блоков демодуляции или декодирования.

Таким образом, итеративная процедура приёма (рис. 4) включает в себя следующие шаги: BCJR демодулятор SEFDM-сигнала, деперемежитель, декодер помехоустойчивого кода и перемежитель. Используя отношения правдоподобия на выходе BCJR демодулятора и априорную информацию, вычисляются логарифмы отношения правдоподобия $L_e(b_k^{(j)}) = L_p(b_k^{(j)}) - L_a(b_k^{(j)})$, которые перемежаются и поступают на вход декодера. Выходом декодера являются либо отношения правдоподобия $L_p^{\text{ПК}}(b_i)$, либо жесткие решения. Отношения правдоподобия $L_p^{\text{ПК}}(b_i)$ используются для расчета априорной информации демодулятора на следующей итерации в виде логарифмов отношения правдоподобия $L_a(b_k^{(j)})$ (представляют собой

перемеженные $L_e^{\text{ПК}}(b_i) = L_p^{\text{ПК}}(b_i) - L_c(b_i)$). На последней итерации выходом декодера являются жесткие решения (биты).

Заметим, что в случае, когда общее количество поднесущих N является степенью числа 2, а L – нет, формирование SEFDM-сигнала на стороне передатчика и формирование N спектральных отсчетов на стороне приёмника возможно с использованием алгоритма быстрого преобразования Фурье с прореживанием по частоте или времени, вычислительная сложность которого равняется $O(\text{Mog}_2 N)$. При этом L -точечные ДФТ и ОДФТ для процесса эквализации могут быть эффективно вычислены с помощью алгоритма Блуштейна, вычислительная сложность которого пропорциональна $O(\text{Mog}_2 M)$, где $M > 2N - 1$ и $M = 2^q$.

Результаты моделирования

Для оценки помехоустойчивости приёма SEFDM-сигналов с параметрами $\alpha = 1/2, 3/4, 7/8$ реализована имитационная модель в среде MATLAB. Количество используемых поднесущих частот $N_n = 200$, общее количество поднесущих частот с учетом защитных интервалов $N = 256$. Количество учитываемых алгоритмом демодуляции коэффициентов $K = 3$. При моделировании использовался помехоустойчивый несистематический сверточный код со скоростью $1/2$, с кодовым ограничением 3 и структурой отводов сдвигового регистра, задаваемой вектором (5, 7) в восьмеричном представлении. В качестве алгоритма декодирования этого кода выступает алгоритм BCJR. Относительный доверительный интервал равняется 0,1, доверительная вероятность 0,99.

В качестве моделей каналов с замираниями использованы многолучевые каналы с замираниями EPA (Extended Pedestrian A model), EVA (Extended Vehicular A model), ETU (Extended Typical Urban model), представленные в стандарте связи LTE. Значения задержек и относительных мощностей лучей для данных каналов приведены в табл. 1–3.

Таблица 1
Значения задержек и относительных мощностей
лучей для канала EPA

Table 1
Path delays and relative power for EPA LTE channel

Задержка луча, нс	Относительная мощность, дБ
0	0,0
30	-1,0
70	-2,0
90	-3,0
110	-8,0
190	-17,2
410	-20,8

Таблица 2
Значения задержек и относительных мощностей
лучей для канала EVA

Table 2
Path delays and relative power for EVA LTE channel

Задержка луча, нс	Относительная мощность, дБ
0	0,0
30	-1,5
150	-1,4
310	-3,6
370	-0,6
710	-9,1
1090	-7,0
1730	-12,0
2510	-16,9

Таблица 3
Значения задержек и относительных мощностей
лучей для канала ETU

Table 3
Path delays and relative power for ETU LTE channel

Задержка луча, нс	Относительная мощность, дБ
0	-1,0
50	-1,0
120	-1,0
200	0,0
230	0,0
500	0,0
1600	-3,0
2300	-5,0
5000	-7,0

Вследствие того, что канал системы связи LTE характеризуется наличием подвижных передатчиков, приёмников и отражателей, на все лучи, представленные в табл. 1–3, оказывает влияние эффект Доплера. В данной работе для моделирования использованы каналы EPA 5Hz, EVA 70Hz, ETU 70Hz с максимальным доплеровским сдвигом частоты 5, 70 и 70 Гц соответственно.

На рис. 5 а приведены кривые помехоустойчивости приёма SEFDM-сигнала с параметром $\alpha = 1/2$ в канале EPA 5Hz. После первой итерации энергетический проигрыш по сравнению с приёмом OFDM-сигнала с тем же помехоустойчивым кодом более 10 дБ с несократимой СВО на бит $\approx 10^{-3}$. После второй итерации энергетический проигрыш уменьшается до 5 дБ, при этом тенденция к наличию несократимой СВО на бит исчезает. Третья итерация позволяет уменьшить энергетический проигрыш алгоритма приёма SEFDM-сигнала по сравнению с приёмом OFDM-сигнала еще на $\approx 0,5$ дБ. Заметим, что для значения $E_b/N_0 = 35$ дБ отношение СВО на бит для SEFDM-сигнала с параметром $\alpha = 1/2$ после первой итерации и для OFDM-сигнала ≈ 100 . После третьей итерации данное отношение уменьшается до значения < 3 .

Рис. 5 б показывает, что для $\alpha = 3/4$ использование второй итерации при приёме SEFDM-сигналов даёт минимальный энергетический выигрыш (менее 0,1 дБ). Таким образом, применение алгоритма итеративного приёма становится нецелесообразным для SEFDM-сигналов с высокими значениями α .

На рис. 5 в представлены кривые помехоустойчивости приёма SEFDM-сигналов с параметрами $\alpha = 1/2, 3/4, 7/8$ для случая $K = 3$ и третьей итерации. Заметим, что для сигнала с параметром $\alpha = 1/2$ демодулятор с параметром $K = 5$ эквивалентен демодулятору с параметром $K = 3$ из-за равенства нулю коэффициентов последовательности $\vec{g} = \{g_k\}_{k=-N/2}^{N/2-1}$ с индексом $k = \pm 2$ [5]. Помехоустойчивость приёма сигнала растёт с увеличением параметра α . Энергетический

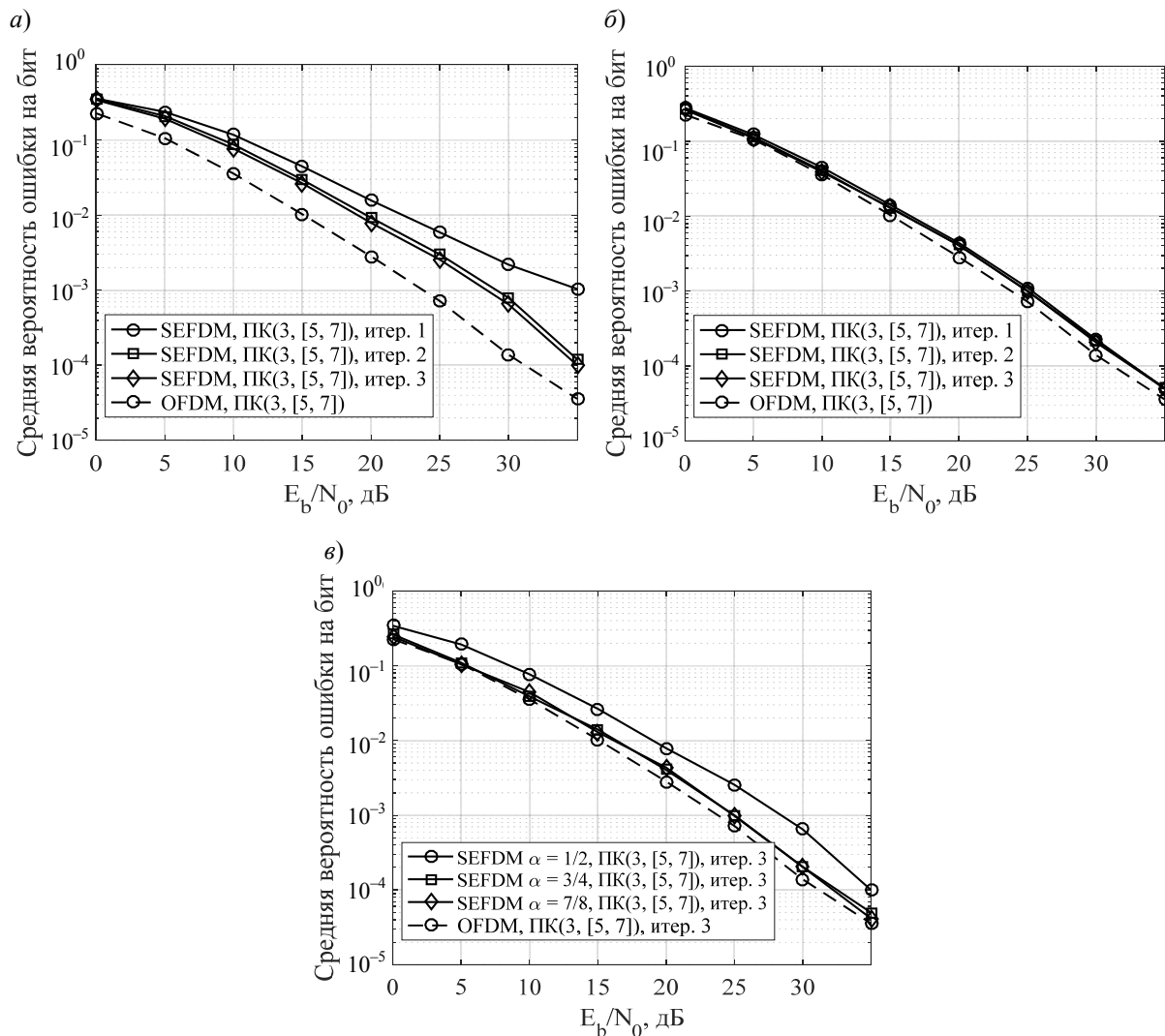


Рис. 5. Характеристики помехоустойчивости SEFDM-сигналов в канале EPA 5Hz, $K = 3$:
а – $\alpha = 1/2$; б – $\alpha = 3/4$; в – $\alpha = 1/2, 3/4, 7/8$, итерация 3

Fig. 5. BER performance of SEFDM-signals in EPA LTE channel, $K = 3$, $\alpha = 1/2, 3/4, 7/8$

проигрыш приёма SEFDM-сигнала с параметром $\alpha = 1/2$ по сравнению с приёмом SEFDM-сигналов с параметром $\alpha = 3/4$ и тем же помехоустойчивым кодом составляет менее 3 дБ для значения СВО на бит $= 10^{-4}$. При этом энергетический выигрыш приёма SEFDM-сигнала с параметром $\alpha = 7/8$ по сравнению с приёмом SEFDM-сигналов с параметром $\alpha = 3/4$ составляет менее 0,5 дБ для значения СВО на бит $= 10^{-4}$.

На рис. 6 представлены характеристики помехоустойчивости SEFDM-сигнала в канале EPA 5Hz. для случая $K = 7$. Для параметров $\alpha = 3/4, 7/8$ и третьей итерации (рис. 6 б, в) энергетический выигрыш от-

носительно случая $K = 3$ составляет менее 0,5 дБ. Для параметра $\alpha = 1/2$ и первой итерации при использовании демодулятора с параметром $K = 7$ (рис. 6 а) энергетический проигрыш по сравнению с помехоустойчивостью приёма OFDM-сигналов с тем же помехоустойчивым кодом уменьшается до 8 дБ в области СВО на бит $\approx 10^{-3}$. При этом для третьей итерации энергетический выигрыш относительно случая $K = 3$ составляет менее 0,5 дБ как и для сигналов с параметром $\alpha = 3/4, 7/8$. Таким образом, эффективность итеративного приёма уменьшается с увеличением параметра K .

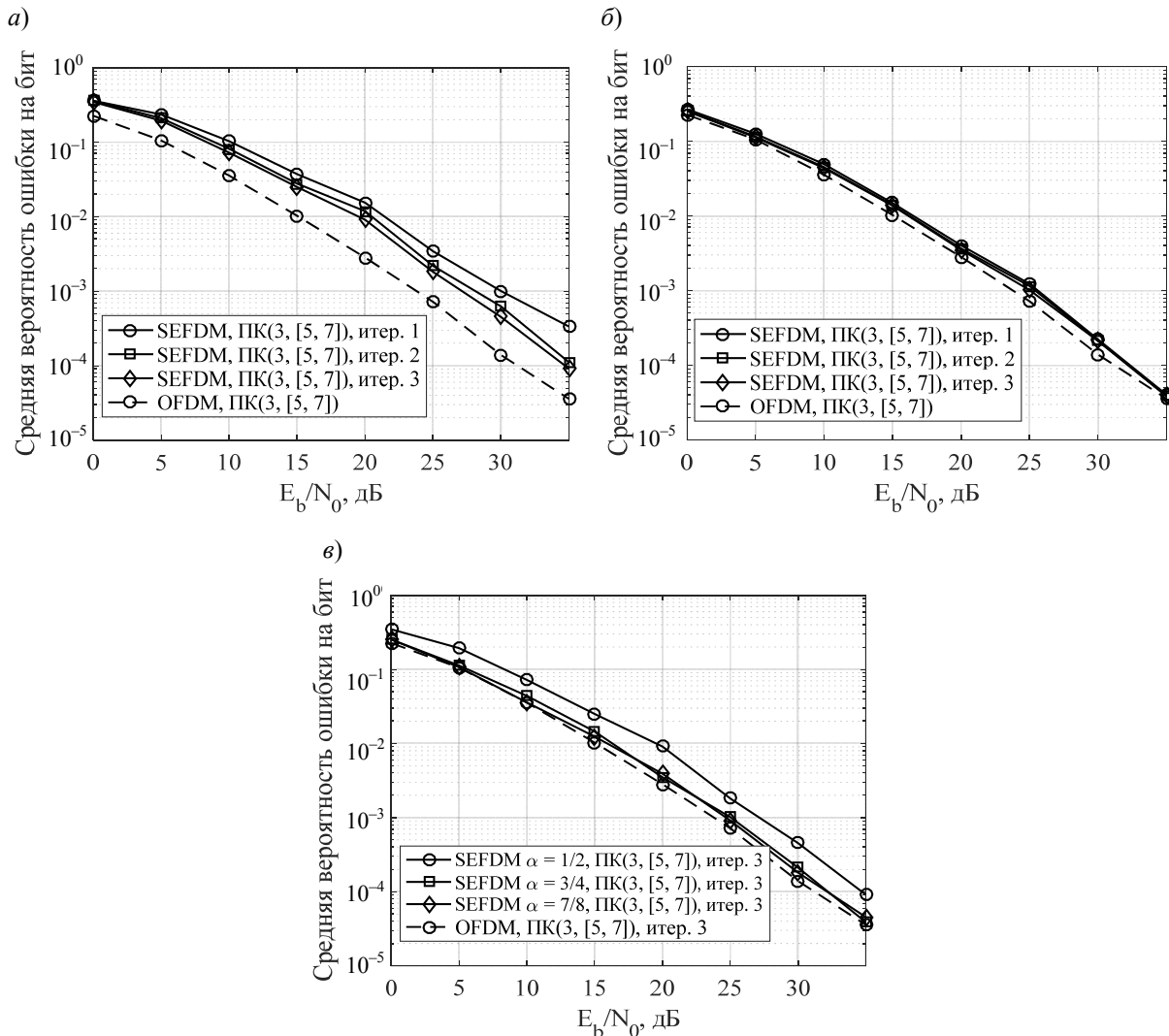


Рис. 6. Характеристики помехоустойчивости SEFDM-сигналов в канале EPA 5Hz, $K = 7$:
 $a - \alpha = 1/2$; $b - \alpha = 3/4$; $c - \alpha = 1/2, 3/4, 7/8$, итерация 3

Fig. 6. BER performance of SEFDM-signals in EPA LTE channel, $K = 7$, $\alpha = 1/2, 3/4, 7/8$

На рис. 7 и 8 представлены характеристики помехоустойчивости SEFDM-сигнала для случая $K = 3$ в каналах EVA 70Hz и ETU 70Hz соответственно. В обоих случаях высокое значение доплеровского сдвига частоты ведет к несократимой СВО на бит для $< 10^{-3}$ при приеме OFDM-сигнала. При этом для SEFDM-сигналов с параметром $\alpha = 1/2$ использование второй итерации позволяет уменьшить уровень несократимой ошибки примерно в два раза с $4 \cdot 10^{-3}$ до $2 \cdot 10^{-3}$ для канала EVA 70Hz (рис. 7 а) и с $6 \cdot 10^{-3}$ до $3 \cdot 10^{-3}$ для канала ETU 70Hz

(рис. 8 а). Для SEFDM-сигналов с параметрами $\alpha = 3/4$ и $\alpha = 7/8$, как и в случае с каналом EPA 5Hz, при использовании второй итерации для каналов EVA 70Hz и ETU 70Hz энергетический выигрыш составляет менее 0,1 дБ (рис. 7 а, б и 8 а, б). При этом энергетический проигрыш относительно приема OFDM-сигнала при значениях сигнал-шум, для которых отсутствует наличие несократимой ошибки, составляет менее 3 дБ при параметре $\alpha < 7/8$ как для канала EVA 70Hz, так и для канала ETU 70Hz.

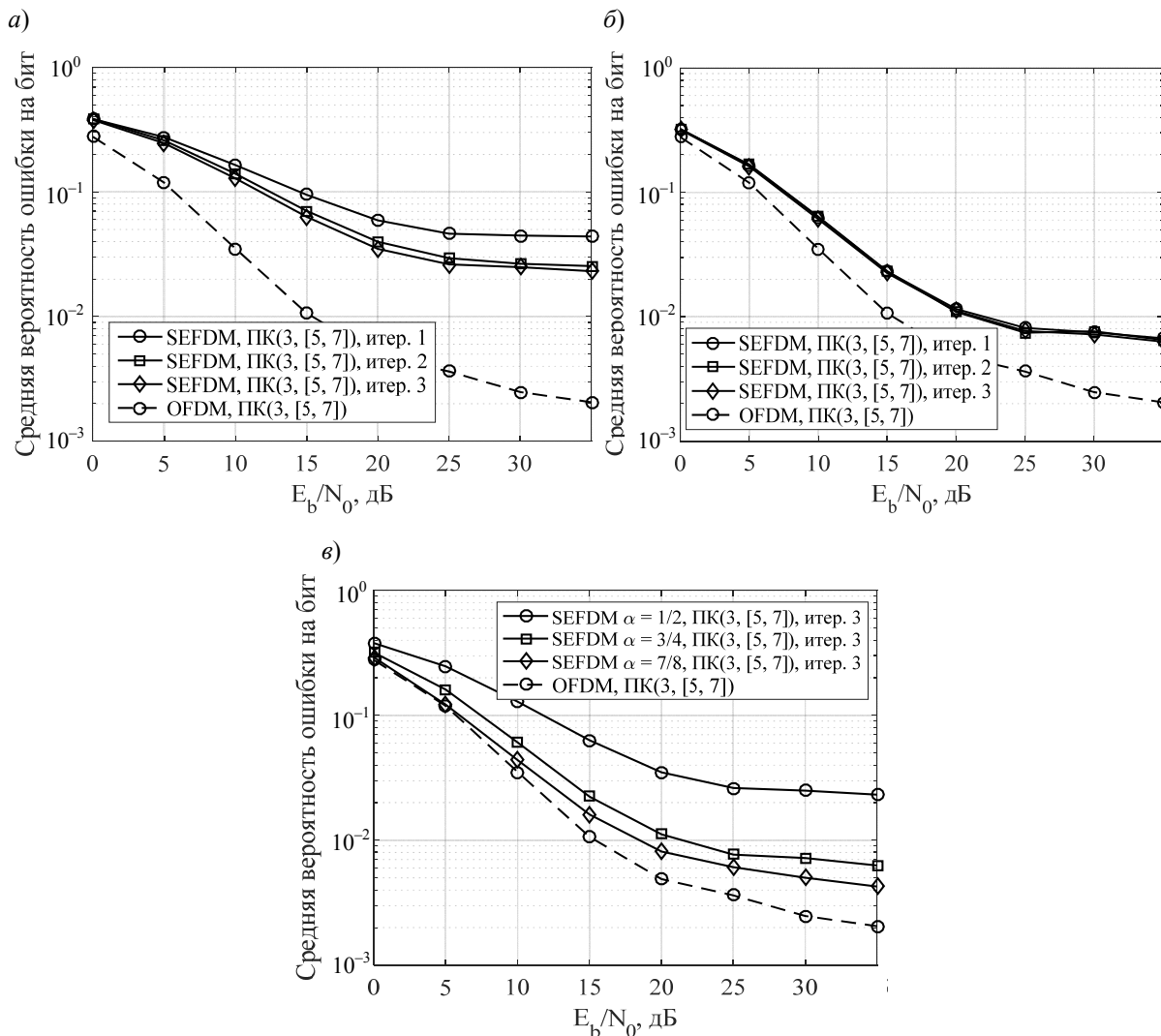


Рис. 7. Характеристики помехоустойчивости SEFDM-сигналов в канале EVA 70Hz, $K = 3$:
а – $\alpha = 1/2$; б – $\alpha = 3/4$; в – $\alpha = 1/2, 3/4, 7/8$, итерация 3

Fig. 7. BER performance of SEFDM-signals in EVA LTE channel, $K = 3$, $\alpha = 1/2, 3/4, 7/8$

Заключение

Проведен анализ эффективности итеративного приёмника SEFDM-сигналов, функционирующего в сочетании с эквалайзером в каналах с частотно-селективными замираниями. Для приёма кодированных SEFDM-сигналов в канале с частотно-селективными замираниями предложен итеративный алгоритм приёма, основанный на совместном применении ZF-эквалайзера эквивалентного OFDM-сигнала, алгоритма BCJR для демодуляции и декодирования сверточного кода. Проведен анализ помехо-

устойчивости приёма SEFDM-сигнала в многолучевых каналах с замираниями EPA 5Hz, EVA 70Hz, ETU 70Hz, который показал, что итеративный метод приёма в случае SEFDM-сигналов с невысокими значениями параметра α уменьшает энергетический проигрыш по сравнению с приёмом OFDM-сигналов до 5 дБ в случае канала EPA 5Hz. При этом для $E_b/N_0 = 35$ дБ отношение СВО на бит для SEFDM-сигнала с параметром $\alpha = 1/2$ и для OFDM-сигнала уменьшается до значения не больше трех. Для каналов EVA 70Hz, ETU 70Hz

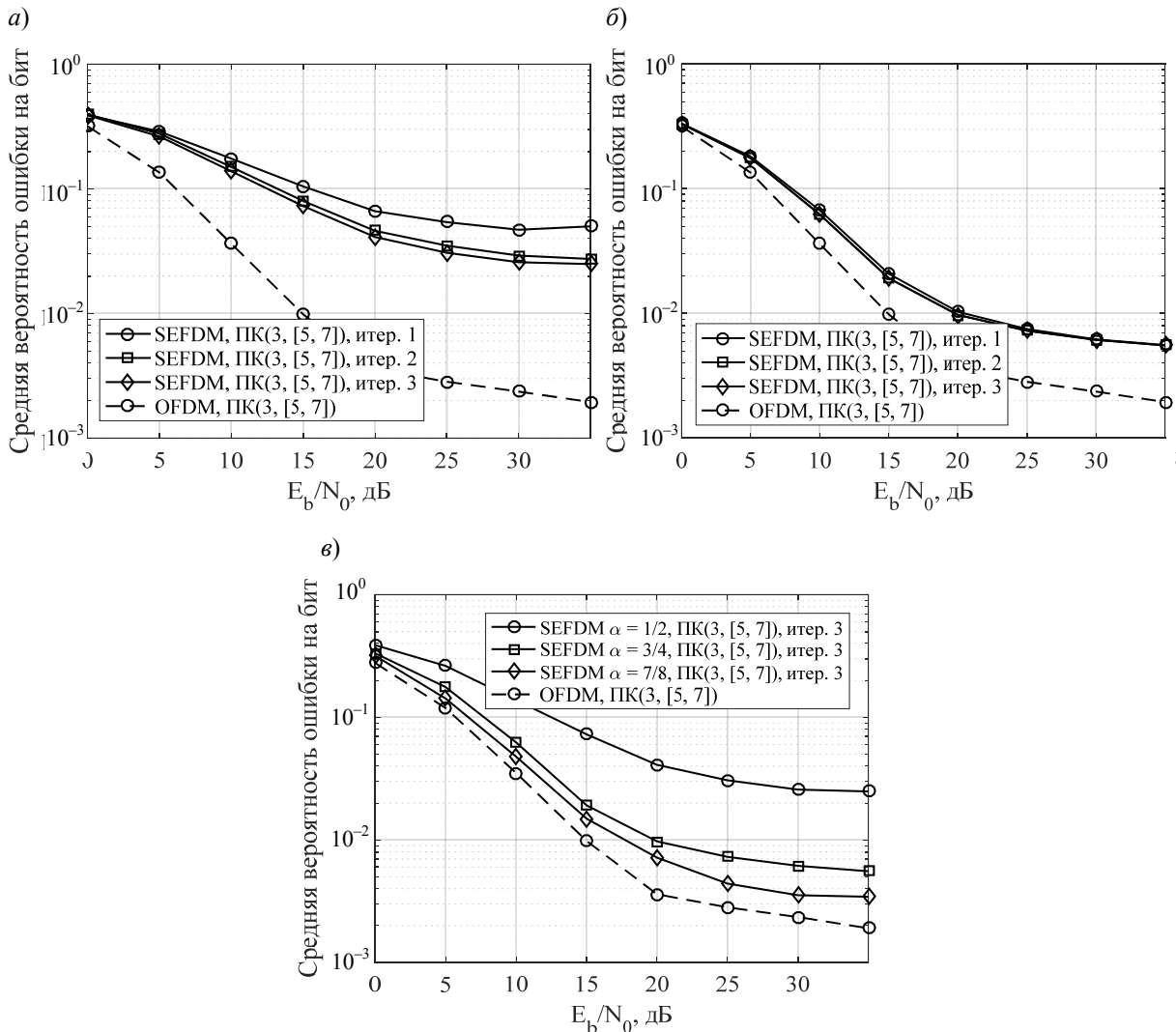


Рис. 8. Характеристики помехоустойчивости SEFDM-сигналов в канале ETU 70Hz, $K = 3$:
 $a - \alpha = 1/2$; $б - \alpha = 3/4$; $в - \alpha = 1/2, 3/4, 7/8$, итерация 3

Fig. 8. BER performance of SEFDM-signals in ETU LTE channel, $K = 3$, $\alpha = 1/2, 3/4, 7/8$

использование итеративного метода приёма в случае SEFDM-сигналов с невысокими значениями параметра α позволяет уменьшить несократимую СВО на бит приблизительно в два раза. При этом, как было показано в [5], в случае канала АБГШ для СВО на бит $= 10^{-4}$ третья итерация позволяет уменьшить энергетический проигрыш по сравнению с приёмом OFDM-сигналов до 1 дБ.

Кроме того, в отличие от работы итеративного приёмника SEFDM-сигнала с тем же сверточным кодом в канале АБГШ [5], для сигналов с высокими значениями па-

раметра α (3/4, 7/8) энергетический выигрыш от использования итеративной схемы приёма в каналах с замираниями составляет менее 0,1 дБ, что делает ее применение нецелесообразным.

Таким образом, в канале с замираниями итеративный приём менее эффективен, чем в канале с АБГШ, что может быть объяснено влиянием процедуры эквализации на спектральную плотность мощности шума и, как следствие, распределением логарифмов правдоподобий канальных битов на входе демодулятора.

СПИСОК ЛИТЕРАТУРЫ

1. Kanaras I., Chorti A., Rodrigues M., Darwazeh I. Spectrally efficient FDM signals: Bandwidth gain at the expense of receiver complexity // Proc. of the IEEE Internat. Conf. on Communications, 2009.
2. Dasalukunte D., Rusek F., Owall V. An iterative decoder for multicarrier faster than-Nyquist signaling systems // Proc. of IEEE Internat. Conf. on Communications. Cape Town, 2010.
3. Kanaras I., Chorti A., Rodrigues M., Darwazeh I. Analysis of suboptimum detection techniques for a bandwidth efficient multi-carrier communication system // Proc. of the Cranfield Multi-Strand Conference. Cranfield University, 2009. Pp. 505–510.
4. Ahmed S., Darwazeh I. Inverse discrete Fourier transform-discrete Fourier transform techniques for generating and receiving spectrally efficient frequency division multiplexing signals // American J. of Engineering and Applied Sciences. 2011. Vol. 4. Pp. 598–606.
5. Rashich A., Kislitsyn A., Fadeev D., Tan Ngoc Nguyen FFT-based trellis receiver for SEFDM signals // Global Communications Conf. 2016. IEEE.
6. Bahl L., Cocke J., Jelinek F., Raviv J. Optimal decoding of linear codes for minimizing symbol error rate // IEEE Transactions on Information Theory. 1974. Vol. IT-20(2). Pp. 284–287.
7. Chorti A., Kanaras I., Rodrigues M.R.D., Darwazeh I. Joint channel equalization and detection of spectrally efficient FDM signals // 2010 IEEE 21st Internat. Symp. on Personal Indoor and Mobile Radio Communications.
8. Isam S., Darwazeh I. Robust channel estimation for spectrally efficient FDM system // Proc. 19th Int. Conf. Telecommun. 2012. Pp. 1–6.
9. Rashich A.V., Gorbunov S.V., Urvantsev A.S. Performance of SEFDM-signals in frequency selective Rayleigh channel // Radiotekhnika. 2018. No. 1. Pp. 57–62.
10. Zavjalov, S.V., Ovsyannikova, A.S., Lavrenyuk, I.I., Volvenko, S.V., Makarov, S.B. Application of optimal finite-length signals for overcoming “Nyquist Limit” // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2018. 11118 LNCS. Pp. 172–180.
11. Pohst M. On the computation of lattice vectors of minimal length, successive minima and reduced basis with applications // ACM SIGSAM Bull. 1981. Vol. 15. Pp. 37–44.
12. Fincke U., Pohst M. Improved methods for calculating vectors of short length in a lattice, including a complexity analysis // Mathematics of Computation. 1985. Vol. 44. Pp. 463–471.
13. Hassibi B., Vikalo H. On the sphere-decoding algorithm I // Expected complexity. IEEE Transactions on Signal Processing. 2005. Vol. 53. No. 8. Pp. 2806–2818.
14. Hassibi B., Vikalo H. On the sphere-decoding algorithm II. Generalizations, Second-order statistics, and applications to communications // IEEE Transactions on Signal Processing. 2005. Vol. 53. No. 8. Pp. 2819–2834.
15. Kislitsyn A., Rashich A., Tan N. Generation of SEFDM-signals using FFT/IFFT // Proc. of the 14th Internat. Conf. NEW2AN 2014 and 7th Conf. ruSMART 2014. 8638 LNCS. Pp. 488–501.
16. Rashich A., Gorbunov S. ZF equalizer and trellis demodulator receiver for SEFDM in fading channels // 26th Internat. Conf. on Telecommunications. 2019. Pp. 1–4.
17. Rashich A., Kislitsyn A., Gorbunov S. Trellis demodulator for pulse shaped OFDM // 2018 IEEE Internat. Black Sea Conf. on Communications and Networking, Batumi, 2018. Pp. 1–4.
18. Robertson P., Hoeher P., Villebrun E. A comparison of optimal and sub-optimal MAP decoding algorithms operating in the logdomain // Proc. of the IEEE Internat. Conf. on Communications. 1999. Vol. 2. Pp. 1009–1013.
19. Bauch G., Franz V. A comparison of soft-in/soft-out algorithms for turbo-detection // Proc. of the Intern. Conf. on Telecomm. 1994.
20. Tschler M., Koetter R., Singer A.C. Turbo equalization: Principles and new results // IEEE Transactions on Communications. 2002. Vol. 50. No. 5.

Статья поступила в редакцию 25.07.2019.

REFERENCES

1. Kanaras I., Chorti A., Rodrigues M., Darwazeh I. Spectrally efficient FDM signals: Bandwidth gain at the expense of receiver complexity. *Proceedings of the IEEE International Conference on Communications*, 2009.
2. Dasalukunte D., Rusek F., Owall V. An iterative decoder for multicarrier faster than-Nyquist signaling systems. *Proceedings of the IEEE International Conference on Communications*, Cape Town, 2010.

3. Kanaras I., Chorti A., Rodrigues M., Darwazeh I. Analysis of suboptimum detection techniques for a bandwidth efficient multi-carrier communication system. *Proceedings of the Cranfield Multi-Strand Conference*. Cranfield University, 2009, Pp. 505–510.
4. Ahmed S., Darwazeh I. Inverse discrete Fourier transform-discrete Fourier transform techniques for generating and receiving spectrally efficient frequency division multiplexing signals. *American Journal of Engineering and Applied Sciences*, 2011, Vol. 4, Pp. 598–606.
5. Rashich A., Kislitsyn A., Fadeev D., Tan Ngoc Nguyen FFT-based trellis receiver for SEFDM signals. *Global Communications Conference 2016. IEEE*.
6. Bahl L., Cocke J., Jelinek F., Raviv J. Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Transactions on Information Theory*, 1974, Vol. IT-20(2), Pp. 284–287.
7. Chorti A., Kanaras I., Rodrigues M.R.D., Darwazeh I. Joint channel equalization and detection of spectrally efficient FDM signals. *2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications*.
8. Isam S., Darwazeh I. Robust channel estimation for spectrally efficient FDM system. *Proceedings of the 19th International Conference on Telecommun.*, 2012, Pp. 1–6.
9. Rashich A.V., Gorbunov S.V., Urvantsev A.S. Performance of SEFDM-signals in frequency selective Rayleigh channel. *Radiotekhnika [Radioengineering]*, 2018, No. 1, Pp. 57–62.
10. Zavjalov, S.V., Ovsyannikova, A.S., Lavrenyuk, I.I., Volvenko, S.V., Makarov, S.B. Application of optimal finite-length signals for overcoming “Nyquist Limit”. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, 11118 LNCS, Pp. 172–180.
11. Pohst M. On the computation of lattice vectors of minimal length, successive minima and reduced basis with applications. *ACM SIGSAM Bull.*, 1981, Vol. 15, Pp. 37–44.
12. Fincke U., Pohst M. Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. *Mathematics of Computation*, 1985, Vol. 44, Pp. 463–471.
13. Hassibi B., Vikalo H. On the sphere-decoding algorithm I. *Expected complexity*. *IEEE Transactions on Signal Processing*, 2005, Vol. 53, No. 8, Pp. 2806–2818.
14. Hassibi B., Vikalo H. On the sphere-decoding algorithm II. Generalizations, Second-order Statistics, and Applications to Communications. *IEEE Transactions on Signal Processing*, 2005, Vol. 53, No. 8, Pp. 2819–2834.
15. Kislitsyn A., Rashich A., Tan N. Generation of SEFDM-signals using FFT/IFFT. *Proceedings of the 14th International Conference NEW2AN 2014 and 7th Conference, ruSMART 2014*, 8638 LNCS, Pp. 488–501.
16. Rashich A., Gorbunov S. ZF equalizer and trellis demodulator receiver for SEFDM in fading channels. *26th International Conference on Telecommunications*, 2019, Pp. 1–4.
17. Rashich A., Kislitsyn A., Gorbunov S. Trellis demodulator for pulse shaped OFDM. *IEEE International Black Sea Conference on Communications and Networking*, Batumi, 2018, Pp. 1–4.
18. Robertson P., Hoeher P., Villebrun E. A comparison of optimal and sub-optimal MAP decoding algorithms operating in the logdomain. *Proceedings of the IEEE International Conference on Communications*, 1999, Vol. 2, Pp. 1009–1013.
19. Bauch G., Franz V. A comparison of soft-in/soft-out algorithms for turbo-detection. *Proceedings of the International Conference Telecomm.* 1994.
20. Tschler M., Koetter R., Singer A.C. Turbo equalization: Principles and new results. *IEEE Transactions on Communications*, 2002, Vol. 50, No. 5.

Received 25.07.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

РАШИЧ Андрей Валерьевич
RASHICH Andrey V.
 E-mail: andrey.rashich@gmail.com

ГОРБУНОВ Сергей Викторович
GORBUNOV Sergey V.
 E-mail: serggv91@gmail.com

DOI: 10.18721/JCSTCS.12406

УДК 621.396.67

МЕТОДИКА ОЦЕНКИ ИНТЕРМОДУЛЯЦИОННЫХ ПОМЕХ ГРУППЫ РАДИОСТАНЦИЙ СРЕДНЕЙ МОЩНОСТИ

А.А. Боробов, Р.Ю. Бородулин

Военная академия связи имени Маршала Советского Союза С.М. Будённого,
Санкт-Петербург, Российская Федерация

Представлена методика оценки интермодуляционных помех, действующих на входе одного из приёмников группы радиостанций средней мощности. Данная методика построена на определении мощностей мешающих сигналов на зажимах приёмной антенны путем вычисления узловых потенциалов. Узловые потенциалы отличаются от обычных тем, что находятся не внутри сегментов, на которые разбивается провод, а в местах соединения проводников, в том числе с нагрузкой. Для расчетов используется вычислительная схема, основанная на методе моментов. В качестве объекта исследования рассмотрена группа радиостанций в составе двух аппаратных. Описан алгоритм расчета мощности интермодуляционной помехи, пересчитанной ко входу приёмника. С помощью данного алгоритма на основе справочных данных произведено численное моделирование помеховой обстановки в районе размещения объекта. Доказано, что представленный способ позволяет упростить задачу расчета мощности в нагрузке приёмной антенны, а также расширить область применения известных методик оценки помеховой обстановки.

Ключевые слова: электромагнитная совместимость, интермодуляционные помехи, развязка антенн, группа радиостанций средней мощности, помеховая обстановка, метод моментов.

Ссылка при цитировании: Боробов А.А., Бородулин Р.Ю. Методика оценки интермодуляционных помех группы радиостанций средней мощности // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 72–83. DOI: 10.18721/JCSTCS.12406

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

METHOD FOR ASSESSING INTERMODULATION INTERFERENCE OF A GROUP OF MEDIUM-POWER RADIO STATIONS

A.A. Borobov, R.Yu. Borodulin

Military Academy of Communications named after Marshal of the Soviet Union S.M. Budenny,
St. Petersburg, Russian Federation

The article presents a methodology for assessing intermodulation interference at the input of one of the receivers of a group of medium-sized radio stations. This technique is based on determining the powers of interfering signals at the terminals of the receiving antenna by calculating the nodal potentials. Nodal potentials differ from ordinary ones in that they are not inside the segments into which the wire is divided, but at the junctions of the conductors, including the load. For calculations, a computational scheme based on the method of moments is used. As an object of study, a group of radio stations consisting of two hardware stations is considered. The authors presented an algorithm for calculating the

power of intermodulation interference, counted to the input of the receiver. Using this algorithm, based on the reference data, a numerical simulation of the interference situation in the area of the object is performed. It is proved that the presented method allows us to simplify the task of calculating the power in the load of the receiving antenna, as well as expand the scope of known methods for evaluating the interference environment.

Keywords: electromagnetic compatibility, intermodulation interference, decoupling of antennas, a group of medium-sized radio stations, interference environment, method of moments.

Citation: Borobov A.A., Borodulin R.Yu. Method for assessing intermodulation interference of a group of medium-power radio stations. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 72–83. DOI: 10.18721/JCSTCS.12406

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Обеспечение электромагнитной совместимости (ЭМС) является одним из основных требований, предъявляемых к узлам связи различного назначения.

Для определения перечня мероприятий, необходимых для обеспечения ЭМС, в каждом конкретном случае требуется качественная оценка помеховой обстановки, сложившейся в районе развертывания узла связи. Данная оценка должна включать анализ потенциальных помех, возникающих из-за нелинейных эффектов в приёмниках. Этот анализ может заключаться не только в определении нежелательных номиналов рабочих частот (для близко расположенных радиостанций), но и в оценке уровней интермодуляционных помех (ИМП) на входе конкретного приёмника.

Цель исследования — разработка методики оценки интермодуляционных помех группы радиостанций средней мощности, работающих на проволочные коротковолновые антенны, расположенные в ближней зоне излучения или вблизи её границы.

Методика расчета узловых потенциалов

В соответствии с рекомендациями союза электросвязи, исходными данными методики расчета ИМП являются уровни мешающих сигналов на входе приёмника [2].

При рассмотрении приёмника (рецептора помех) с подключенной к нему проволочной приёмной антенной и мешающего

передатчика (эмиттера помех) с подключенной к нему проволочной передающей антенной, значения ИМП можно рассчитать, например, путём решения системы интегральных уравнений методом моментов (ММ). Данная система формулируется в частотном представлении и основывается на методе комплексных амплитуд [10]. При составлении подобных вычислительных схем, в том числе и при использовании известных программ, значения токов можно получить либо на поверхности, либо в узлах соединения проводов, в том числе на зажимах нагруженной приёмной антенны [11].

При расчете мощности принимаемого сигнала, выделяемой на нагрузке (приёмнике), возникают существенные трудности вычислительного характера, вызванные многократным ручным перезапуском программ с различным набором исходных данных, что может привести к накоплению различного рода ошибок. Так, для нахождения тока в нагрузке (имитации приёмника) потребуется проведение ряда манипуляций, таких как нахождение ЭДС на разомкнутых зажимах антенны (без нагрузки) и вычисление входного сопротивления приёмной антенны в режиме передачи.

Существенно упростить задачу расчета мощности в нагрузке можно путём модернизации известных вычислительных алгоритмов с целью нахождения так называемых узловых потенциалов. Узловые потенциалы отличаются от обычных тем, что находятся

не внутри сегментов, на которые разбивается провод, а в местах соединения проводников, в том числе с нагрузкой.

Рассмотрим процедуру нахождения узловых потенциалов на примере известной вычислительной схемы, основанной на ММ с кусочно-постоянными базисными функциями.

Одним из типов интегральных уравнений, решаемых методом моментов, является интегральное уравнение для электрического поля (ИУЭП), которое для тонкопроволочных структур может иметь вид уравнения Харрингтона [12]:

$$E_z^i(0, z) = \int_{-L/2}^{L/2} \left[i\omega\mu I(z') - \frac{1}{i\omega\epsilon} \frac{\partial I(z')}{\partial z'} \frac{\partial}{\partial z} \right] \times \times \frac{\exp(-ikr)}{4\pi r} dz', \quad (1)$$

где L — длина провода.

Здесь неизвестными являются значения токов $I(z')$ в выбранных точках z проводов приёмной и передающей антенн. Искомое электромагнитное поле (ЭМП) — это касательные к проводникам составляющие вектора электрического и магнитного полей $\mathbf{E}_t$, $\mathbf{H}_t$, что есть сумма рассеянного на проводниках поля (поля, создаваемого токами, текущими по проводникам) $\mathbf{E}$, $\mathbf{H}$ и заданного действующего поля (поля источника) $\mathbf{E}_0$, $\mathbf{H}_0$.

Выражение для скалярного потенциала $V(r)$ через решение уравнений Гельмгольца имеет вид [1]:

$$V(r) = -\frac{1}{i\omega\epsilon} \int \mathbf{j}^{ct}(\mathbf{r}') \cdot \text{grad}'(G(|\mathbf{r} - \mathbf{r}'|)) d\mathbf{r}'^3 = = \frac{1}{i\omega\epsilon} \text{div} \int G(|\mathbf{r} - \mathbf{r}'|) \mathbf{j}^{ct}(\mathbf{r}') \cdot d\mathbf{r}'^3, \quad (2)$$

где $G(|\mathbf{r} - \mathbf{r}'|) = G(R) = \frac{e^{ikR}}{4\pi R}$ — функция Грина.

Выражение для нахождения потенциала V рассеянного поля, наводимого касательной составляющей к поверхности провода вектора напряженности электрического поля E падающей волны, можно представить в виде [13]:

$$V^{(n)}(s) = \frac{1}{i\omega\epsilon} \sum_{m=1}^{N_w} \int_0^l G(R^{(n,m)}(s, s')) \frac{d}{ds'} I^{(m)}(s') ds', \quad (3)$$

где s, s' — расстояния вдоль n -го и m -го сегментов разбиения, отсчитываемые от их начала, соответственно; N_w — общее число сегментов проволочной структуры; l — длина сегмента; $G(R)$ — функция Грина; $R^{(n,m)}$ — расстояние между n -м и m -м сегментами; $I^{(m)}$ — комплексная амплитуда тока m -го сегмента на расстоянии s' от его начала.

Алгоритм нахождения потенциала на n -м сегменте в его точках s (во всех точках, кроме узловых) включает следующие шаги.

1. Найти токи в точках s' m -го сегмента путем решения системы линейных алгебраических уравнений, составленной из уравнения Харрингтона.

2. Найти производные токов на m -м сегменте в точках s' путём разбиения сегментов на отрезки, воспользовавшись разностной схемой.

3. Вычислить функцию Грина между точками s, s' n -го и m -го сегментов.

4. Численно проинтегрировать вдоль сегмента найденные величины.

5. Повторить шаги 1–4 для каждого сегмента проволочной структуры и просуммировать полученные значения интегралов с учётом первого множителя.

Далее находятся узловые потенциалы. Для этих целей можно использовать два типа граничных условий:

- условие, основанное на выполнении закона Кирхгоффа для токов проводников, сходящихся в узел (сумма токов равна нулю);
- условие, основанное на выполнении условия непрерывности зарядов (потенциалов) при переходе от сегмента к сегменту, на которые разбивается анализируемый проводник.

Для метода моментов с кусочно-постоянным базисом идея соблюдения непрерывности потенциала заключается в том, что ставится задача выполнения не только условия равенства производных токов в узлах, но и условия равенства потенциалов в узлах соединения проводов. Это требуется для исключения их «скачков» при переходе через узлы, соединяющие соседние сегменты проводов.

Можно экстраполировать значения потенциалов с отрезков сетки разбиения сегментов, граничащих с узлом, в данный узел, согласно следующим выражениям:

$$U = V^{(n)}(h_n) - h_n \frac{d}{ds} V^{(n)}(h_n) \quad (4 \text{ а})$$

или

$$U = V^{(n)}(h_n) + h_n \frac{d}{ds} V^{(n)}(l_n - h_n), \quad (4 \text{ б})$$

где U — значение потенциала в узле; h_n — шаг сетки на n -м проводе.

Выражение (4 а) используется в том случае, когда начало n -го провода подключено к узлу. Выражение (4 б) — когда подключен конец.

Таким образом, требование непрерывности зарядов заключается в том, что потенциалы, экстраполированные в узел, становятся равны одной и той же постоянной величине U . В случае если узел является свободным (например, конец проволочной антенны), потенциал V в свободном пространстве определяется согласно выражению (3).

После применения выражений (3) и (4) средняя мощность может быть получена через ток в первом и втором узлах, подключенных к нагрузке I_{n1} , I_{n2} , и потенциалы данных узлов U_{n1} , U_{n2} , соответственно, как

$$P_n = 0,5 \operatorname{Re}(I_{n1} U_{n1}^* + I_{n2} U_{n2}^*). \quad (5)$$

Определив мощности мешающих сигналов на входе приёмника, можно с высокой точностью рассчитать и уровень ИМП, создаваемой ими, воспользовавшись рекомендациями союза электросвязи [2].

Алгоритм оценки интермодуляционных помех

Алгоритм оценки интермодуляционных помех (рис. 1) может включать в себя определение номиналов потенциально опасных частот передачи близко расположенной радиостанции и мощностей сигналов на данных частотах. Далее по методике [2] можно рассчитать уровни ИМП, пересчитанные ко входу приёмника.

В качестве исходных данных алгоритма вводятся следующие показатели:

$f_{\text{пер}}, f_{\text{пр}}$ — частоты передачи и приёма;

x_1, y_1, z_1 — координаты точки установки мачты передающей антенны;

x_2, y_2, z_2 — координаты точки установки мачты приёмной антенны;

h_1, h_2 — высота подъёма передающей и приёмной антенн соответственно;

$\varepsilon_1, \mu_1, \sigma_1$ — электрические параметры воздуха;

$\varepsilon_2, \mu_2, \sigma_2$ — электрические параметры почвы;

L_1 — расстояние между передающей и приёмной антеннами собственной радиостанции;

P_s — уровень полезного сигнала на входе приёмника;

$P_{\text{ш}}$ — уровень шума на входе приёмника;

A — требуемое превышение уровня сигнала над уровнем помехи.

После ввода исходных данных счётчикам потенциально опасных частот соседнего передатчика присваиваются нулевые значения (блок 2). В блоках 3–14 производится вычисление и запись всех возможных номиналов частот соседнего передатчика, способных сформировать ИМП m -го и k -го порядков. Далее в блоках 15–17 производится расчёт развязки B между антеннами эмиттера и рецептора помех. Для расчёта развязки на вход передающей антенны подаётся сигнал генератора постоянного тока с уровнем P_i на исследуемой частоте f_k , затем по методике, описанной выше, рассчитывается выделяемая на нагрузке приёмной антенны мощность P_r . Развязка B определяется как отношение [8]:

$$B = 10 \lg \left(\frac{P_i(f_k)}{P_r(f_k)} \right). \quad (6)$$

Далее вычисляется мощность мешающего сигнала на зажимах приёмной антенны на всех определенных частотах простым вычитанием величины B из номинальной мощности передатчика $P_{\text{пер}}$ (блок 18). Имея значения данных мощностей, можно вычислить уровни ИМП по любой из известных методик (блоки 19–21).

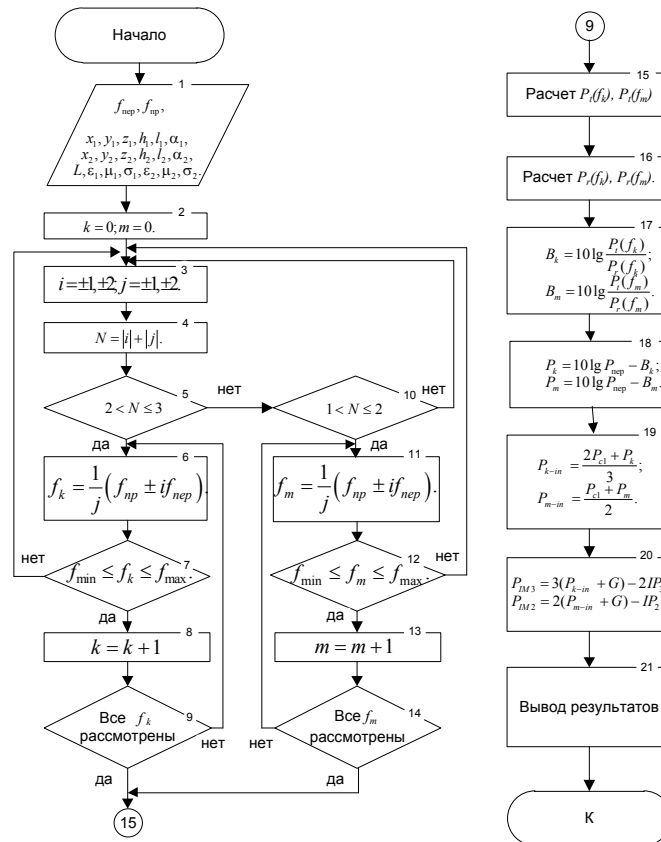


Рис. 1. Алгоритм расчета уровня ИМП на входе приёмника
Fig. 1. The algorithm for calculating the level of IPM at the input of the receiver

Представленный способ определения мощности помехи на входе приёмника впервые позволяет с высокой точностью рассчитывать развязку между близкорасположенными проволочными коротковолновыми антеннами, произвольно ориентированными в пространстве. Кроме того, он учитывает влияние подстилающей поверхности и металлических конструкций (антенных мачт) на токи, наводимые в проводниках.

Сравним результаты анализа развязки антенн по методике расчёта узловых потенциалов с результатами, полученными другими известными способами. Один из них основан на вычислении взаимного сопротивления антенн:

$$S_{nm}(f) = 10 \lg \left\{ \frac{1}{2} \frac{\left[\operatorname{Re} \left(\frac{Z_{12}^2}{Z_{11}^2} \right) \right]^2 + \left[\operatorname{Im} \left(\frac{Z_{12}^2}{Z_{11}^2} \right) \right]^2}{2 - \left[\operatorname{Re} \left(\frac{Z_{12}^2}{Z_{11}^2} \right) \right]^2 + \left[\operatorname{Im} \left(\frac{Z_{12}^2}{Z_{11}^2} \right) \right]^2} \right\}, \quad (7)$$

где Z_{11} – собственное сопротивление передающей антенны; Z_{12} – взаимное сопротивление передающей и приёмной антенн.

Величину взаимного сопротивления Z_{12} можно определить различными способами, например, методом наводимых ЭДС, подробно описанном в [9].

Достоинством данного метода является достаточно высокая точность результатов и возможность расчета развязки антенн в ближней зоне. Существенными недостатками являются критические ограничения на взаимное размещение, ориентацию и форму антенн (симметричные проволочные вибраторы в параллельных плоскостях). Кроме того, при определении взаимных сопротивлений методом наводимых ЭДС не учитываются параметры подстилающей поверхности и влияние мачты приёмной антенны.

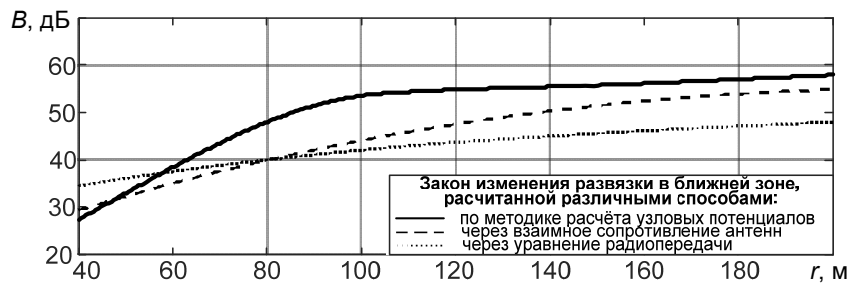


Рис. 2. Сравнение способов определения развязки антенн
Fig. 2. Comparison of methods for determining the isolation of antennas

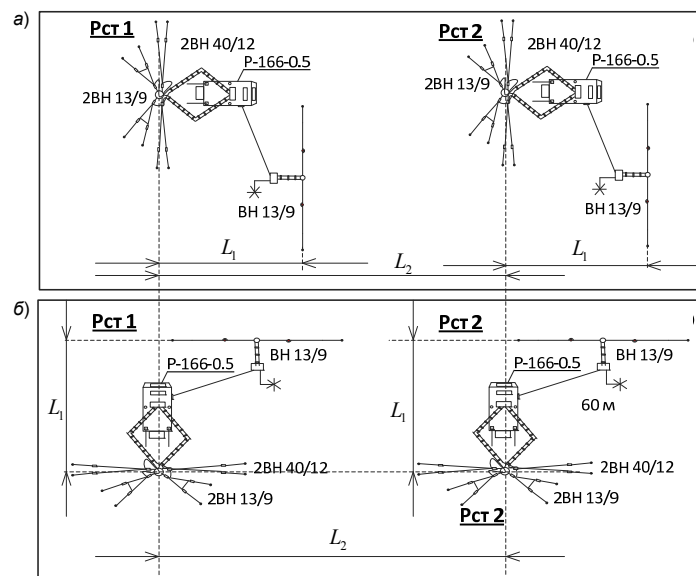


Рис. 3. Группа радиостанций средней мощности (вариант)
Fig. 3. A group of medium-sized radio stations (option)

При рассмотрении помехи с позиции полезного сигнала её мощность на входе рецептора оценивается на основе уравнения радиопередачи [18]. При таком подходе определяющим фактором является уровень потерь сигнала на трассе распространения электромагнитной волны от эмиттера к рецептору. К сожалению, существующие модели расчёта данного параметра, подробно описанные в [14, 15], неприменимы либо для декаметрового диапазона волн, либо для ближней зоны излучения антенн.

Сравнение результатов расчёта развязки различными способами в зависимости от расстояния между антеннами (рис. 1) показывает, что способ анализа, основанный на расчёте узловых потенциалов, даёт более

объективную картину помеховой обстановки за счёт учёта влияния земли и мачты антенны.

Применим описанный выше алгоритм для расчёта уровней потенциальных ИМП 2-го и 3-го порядка на входе произвольного приёмника КВ диапазона группы радиостанций средней мощности, состоящей из двух аппаратных (рис. 2).

При одновременной работе передатчиков обеих станций на приёмник одной из них действуют три сигнала, один из которых S_0 является полезным, а два других — мешающими. Они исходят от собственного передатчика S_{m1} и передатчика соседней радиостанции S_{m2} . Определение мощности полезного сигнала производится по методам, подробно описанным в [4–6].

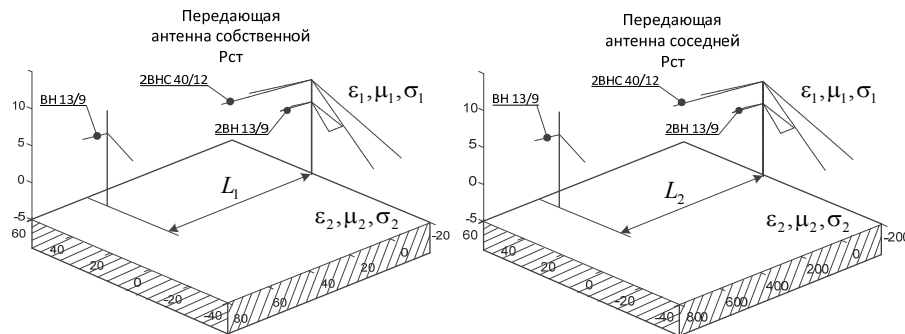


Рис. 4. Модель антенного поля группы радиостанций средней мощности
Fig. 4. Model of the antenna field of a group of medium-sized radio stations

Для расчета мощностей сигналов S_{m1} и S_{m2} необходимо отдельно рассмотреть дуэльные ситуации взаимодействия приёмника (рецептора помех) с каждым из мешающих передатчиков (эмиттерами помех).

Постановка задачи

Пусть имеются две близко расположенные радиостанции средней мощности Рст1 и Рст2, работающие в разных радионаправлениях (см. рис. 2). Район развёртывания – Санкт-Петербург. Радиостанция Рст1 работает в радионаправлении с корреспондентом, находящимся в районе г. Мурманск (расстояние между корреспондентами около 1000 км). Частота передачи $f_{\text{пер}} = 12\,200$ кГц, частота приёма $f_{\text{пр}} = 12\,000$ кГц. Вид работы – однополосная модуляция (ОМ) с подавленной несущей (J3E)¹. Минимально необходимый уровень полезного сигнала на входе приёмника $P_s = -76$ дБм. Антенна, работающая на передачу – 2ВН 13/9, приёмная антенна – ВН13/9². Почва в районе развёртывания радиостанций влажная ($\epsilon_2 = 10$, $\mu_2 = 1$, $\sigma_2 = 10^{-2}$).

Расстояние между передающей и приёмной антеннами радиостанции Рст1 $L_1 = 60$ м. Дистанция между передающей антенной Рст2 и приёмной антенной Рст1 (L_2) выбирается поочередно равной 200, 300, 400 и 500 м.

Необходимо рассчитать уровень потенциальных интермодуляционных помех,

создаваемых на входе приёмника радиостанции Рст1 (рецептора помех) передатчиками радиостанций Рст1 и Рст2 (эмиттерами помех), и сделать вывод о степени выполнения условия ЭМС при различных вариантах настройки передатчика Рст2.

Для вычисления номиналов частот соседнего передатчика $f_{\text{пер}2}$, способных вызывать ИМП в приёмнике радиостанции Рст1, воспользуемся известным условием возникновения интермодуляционных составляющих [16]:

$$nf_1 \pm mf_2 = f_3, \quad (8)$$

где f_1, f_2 – частоты мешающих сигналов; f_3 – частота приёма полезного сигнала; $n, m = \pm 1, \pm 2, \pm 3, \dots$

Из уравнения (8) получим:

$$f_2 = \frac{1}{m}(f_3 \pm nf_1). \quad (9)$$

Решая уравнение (9) с учетом условия

$$N = |n| + |m| \leq 3, \quad (10)$$

находим возможные значения $f_{\text{пер}2}$: 36400, 24200, 12100, 12400 кГц.

Далее по описанному выше алгоритму (рис. 1) рассчитаем уровни мощности мешающих сигналов на входе приёмника.

Численное моделирование помеховой обстановки группы радиостанций

Проведём моделирование работы передающей и приёмной антенн, развёрнутых на модели местности (рис. 4).

¹ Изделие Р–170П. Техническое описание. ЦЛ2.003.143 ТО.

² Передатчик Артек-1-КВ. Техническое описание. ШКИС.464124.002.



Таблица 1

Уровни мощности мешающих сигналов на входе приёмника

Table 1

Power levels of interfering signals at the receiver input

$f_{\text{пр}},$ кГц	Параметры мешающего сигнала собственного передатчика		Параметры мешающего сигнала соседнего передатчика		$L_1, \text{ м}$	$L_2, \text{ м}$
	$f_{\text{пер1}}, \text{ кГц}$	$P, \text{ дБм}$	$f_{\text{пер2}}, \text{ кГц}$	$P, \text{ дБм}$		
12000	12200	4,5	36400	−55,5	60	200
		4,5	24200	−70,5		
		4,5	12100	−3		
		4,5	12400	−3		
12000	12200	4,5	36400	−50	60	300
		4,5	24200	−59		
		4,5	12100	−9		
		4,5	12400	−9,5		
12000	12200	4,5	36400	−52,5	60	400
		4,5	24200	−65		
		4,5	12100	−14		
		4,5	12400	−14		
12000	12200	4,5	36400	−55	60	500
		4,5	24200	70,5		
		4,5	12100	−17		
		4,5	12400	−18		

В результате расчетов методом моментов получаем значения мощности P (дБм) мешающих сигналов S_{m1} и S_{m2} на частотах $f_{\text{пер}2}$ (табл. 1).

Теперь по методике [2] вычислим мощности потенциальных интермодуляционных помех, пересчитанные ко входу приёмника.

На графике (рис. 5) пунктирной горизонтальной линией обозначен минимально допустимый уровень сигнала для данной радиотрассы, при котором выполняется требуемое превышение уровня сигнала над уровнем помехи.

Как видно из результатов расчёта, на некоторых частотах приёмник оказывается поражённым (уровень помехи превышает уровень сигнала), и требование по ЭМС не выполняется.

При повороте передающей и приёмной антенн на угол 90° (рис. 3 б), уровень ИМП, в целом, снижается (рис 5 б). Таким образом, из графика видно, что эффект от изменения угла направленности антенн соизмерим с эффектом от территориального разнеса в ближней зоне.

Согласно исследованиям помеховой обстановки на территории Ленинградской области, число и уровень мешающих сигналов описывались данными, представленными в табл. 2 [7].

В качестве примера проведем моделирование работы радиостанции Рст1 в помеховой обстановке, описанной в табл. 2. Для упрощения исследования ограничимся анализом интермодуляционных продуктов третьего порядка.

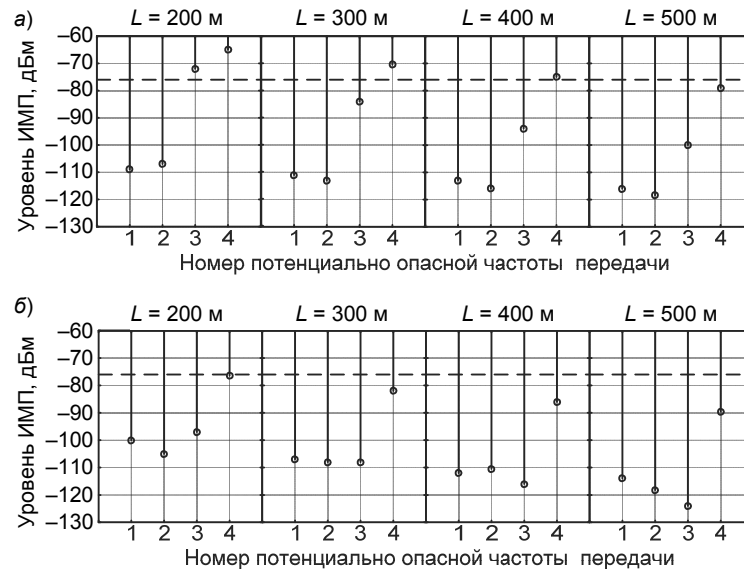


Рис. 5. Уровни мощности ИМП при различных вариантах взаимного расположения антенн
Fig. 5. The power levels of the IMP with various options for the relative positioning of the antennas

Таблица 2

Данные помеховой обстановки в диапазоне частот 12000...13000 кГц на территории Ленинградской области

Table 2

Data interference environment in the frequency range 12000...13000 kHz in the Leningrad Region

Диапазон частот, кГц	Число мешающих сигналов при уровнях, дБ/мкВ (дБ относительно 1 мкВ)					
	40...50	50...60	60...70	70...80	80...90	90...100
12000...13000	30	20	10	5	2	2

В низкочастотной области диапазона значительный вклад в общий уровень помех вносят промышленные помехи, которые, складываясь на входе приёмного устройства, образуют примерно равный по спектру шум, статистические свойства которого достаточно точно описываются стационарным случайным процессом с нормальным распределением вероятностей мгновенных значений амплитуд [17]:

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{\xi^2}{2}} d\xi, \quad (11)$$

где $u = \bar{z} - z_{\text{тр}} / \sigma_z$ — аргумент функции; $\bar{z} = P_c - P_n$, дБ — отношение медианных значений мощности сигнала к мощности помехи; $\sigma_z = \sqrt{\sigma_c^2 + \sigma_n^2}$ — среднеквадратическое отклонение величины z ; σ_c^2, σ_n^2 — дисперсия уровней сигналов и помех соответственно.

Очевидно, что при формировании интермодуляционных продуктов в приёмнике наибольший вклад в мощность помехи вносит передатчик собственной радиостанции. Здесь важен номер гармоники излучения, поскольку уровень гармоник значительно ослаблен по сравнению с полезным сигналом (как правило, уровень высших гармонических составляющих сигнала не превышает -65 дБ).

Подставим в алгоритм (см. рис. 1) вместо мощности мешающих сигналов соседнего передатчика $P_k(P_m)$ уровни помех из табл. 2, значения которых в указанных диапазонах распределены по закону (11). Тогда, предполагая, что в формировании ИМП участвует первая гармоника сигнала собственного передатчика, получим около 20 возможных помех с мощностью выше допустимой (рис. 6).



Рис. 6. Уровни мощности возможных ИМП третьего порядка с учётом помеховой обстановки в районе расположения рецептора

Fig. 6. Power levels of possible third-order UTIs taking into account the interference environment in the region of the receptor location

Учитывая существенный прирост количества РЭС (по сравнению с концом XX века) и малую область анализируемого диапазона частот, можно сделать вывод о неблагоприятной помеховой обстановке, в которой будет находиться исследуемый приёмник в современных условиях.

Для борьбы с интермодуляцией наиболее эффективными и целесообразными мерами являются те, которые направлены на снижение влияния излучения собственного передатчика при формировании ИМП. Такими мерами могут быть частотный, территориальный и поляризационный разнос эмиттера и рецептора помех. К сожалению, манёвр частотами при современном уровне освоения диапазона затруднён, а территориальный разнос ограничен возможностями кабельного оборудования.

Наиболее перспективным направлением является улучшение пространственной селективности антенн за счет изменения их конфигурации с минимальным усложнением конструкции. При этом основным объектом оптимизации должна выступать приёмная антенна, поскольку изменение направленных свойств передающей антен-

ны повлияет на уровень сигнала на входе приёмника корреспондента и, как следствие, может снизиться качество связи.

Улучшение электрических характеристик приёмной антенны в определённых направлениях прихода волны позволит изменять её ориентацию относительно собственной передающей антенны. Такая мера позволит снижать уровень помех от близко расположенного передатчика, избегая при этом недопустимых потерь мощности полезного сигнала на входе приёмника, обусловленных несовпадением максимума характеристики направленности приёмной антенны.

Заключение

Представленная методика оценки ИМП на входе приёмника позволяет с большой точностью определять мощность мешающих сигналов на зажимах проводочной антенны. Она даёт возможность производить расчеты для ближнего поля излучения, что расширяет область применения некоторых известных методик оценки ИМП (например, [3]) и впервые позволяет применить их для декаметрового диапазона радиоволн.

СПИСОК ЛИТЕРАТУРЫ

1. Бородулин Р.Ю. Численные методы электродинамики: Монография. СПб.: ВАС, 2016. 180 с.
2. Рекомендация Международного союза электросвязи МСЭ-R SM.1134-1. Расчет интермодуляционных помех в сухопутной подвижной службе.
3. ГОСТ Р 55898–2013. М.: Стандартинформ, 2014.
4. Серков В.П. Распространение радиоволн и антенные устройства. СПб.: ВАС, 1981. С. 1–468.
5. Мешалкин В.А., Сосунов Б.В. Основы энергетического расчёта радиоканалов. СПб.: ВАС, 1991. С. 110.

6. Прохоров В.К., Шаров А.Н. Расчёт показателей эффективности радиосвязи. СПб.: ВАС, 1982. С. 1–132.
7. Комарович В.Ф., Сосунов В.Н. Случайные радиопомехи и надёжность КВ связи. М.: Связь, 1977. 136 с.
8. Вахлаков В.Р., Рожков А.Г., Сосунов Б.В., Чернолес В.П., Школин Ю.Д., Ярошенко В.С. Основы обеспечения электромагнитной совместимости радиоэлектронных средств. СПб.: ВАС, 1991. 206 с.
9. Айзенберг Г.З. Коротковолновые антенны. М.: Гос. изд-во литературы по вопросам связи и радио, 1962. 815 с.
10. Бородулин Р.Ю. Применение концептуальных моделей численных методов электродинамики для анализа характеристик вибраторов в бесконечных диссипативных средах // Информатика. Телекоммуникации. Управление. Научно-технические ведомости СПбГПУ. 2016. № 4 (236). С. 29–42. DOI: 10.5862/JCSTCS.252.3
11. Вычислительные методы в электродинамике. Под ред. Р. Миттра. М.: Мир, 1977. 243 с.
12. Richmond J. Computer program fir Thin-Wire Structures in Homogeneous Conducting Medium // National Technical Information Service. Springfield: VA, NASA contractor Rep., July 1973. CR-2399.
13. Burke G.J., Poggio A.J. Numerical Electromagnetic Code (NEC) method of moments. Lawrence Livermore Laboratory, 1981. 81 p.
14. ITU-R Rec. P.368-7. Ground-wave propagation curves for frequencies between 10 kHz and 30 MHz. ITU, 1992.
15. ITU-R Rec. P.1546-1. Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz. ITU, 2003.
16. Перцов С.В., Шуцкой К.А. Входные цепи радиоприёмников. М.: Энергия, 1973.
17. Головин О.В. Декаметровая связь. М.: Радио и связь, 1990. 240 с.
28. Бородулин Р.Ю., Боробов А.А. Методика анализа развязки проволочных антенн методом моментов с использованием выполнения условий непрерывности зарядов // Информация и космос. 2019. № 3. С. 18–24.

Статья поступила в редакцию 11.09.2019.

REFERENCES

1. Borodulin R.Yu. *Numerical methods of electrodynamics*: Monograph. St. Petersburg: Military Academy of Communications Publ., 2016. 180 p. (rus)
2. ITU-R Rec. P.SM.1134-1. *Calculation of intermodulation interference in the land mobile service*. (rus)
3. GOST R 55898–2013. Moscow: Standartinform, 2014. (rus)
4. Serkov V.P. *Radio wave propagation and antenna devices*. St. Petersburg: Military Academy of Communications Publ., 1981. Pp. 1–468. (rus)
5. Meshalkin V.A., Sosunov B.V. *Fundamentals of energy calculation of radio channels*. St. Petersburg: Military Academy of Communications Publ., 1991. P. 110. (rus)
6. Prokhorov V.K., Sharov A.N. *Calculation of indicators of the effectiveness of radio communications*. St. Petersburg: Military Academy of Communications Publ., 1982. Pp. 1–132. (rus)
7. Komarovich V.F., Sosunov V.N. *Random radio interference and HF communications reliability*. Moscow: Radio i Svyaz Publ., 1977. 136 p. (rus)
8. Vakhlakov V.R., Rozhkov A.G., Sosunov B.V., Chernoles V.P., Shkolin Yu.D., Yaroshenko V.S. *Fundamentals of ensuring electromagnetic compatibility of electronic equipment*. St. Petersburg: Military Academy of Communications Publ., 1991. 206 p. (rus)
9. Eisenberg G.Z. *Shortwave antennas*. Moscow: Gos. izd-vo literatury po voprosam svyazi i radio Publ., 1962. 815 p. (rus)
10. Borodulin R.Yu. Using of numerical electrodynamical methods for analyzing the characteristics of vibrators in infinite dissipative media. *St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems*, 2016, No. 4 (236), Pp. 29–42. DOI: 10.5862/JCSTCS.252.3. (rus)
11. *Computational methods in electrodynamics*. Ed. R. Mittra. Moscow: Mir Publ., 1977. 243 p. (rus)
12. Richmond J. *Computer program fir thin-wire structures in homogeneous conducting medium*. National Technical Information Service. Springfield: VA, NASA Contractor Rep., July 1973. CR-2399.
13. Burke G.J., Poggio A.J. *Numerical electromagnetic code (NEC) method of moments*. Lawrence Livermore Laboratory, 1981. 81 p.
14. ITU-R Rec. P.368-7. *Ground-wave propagation curves for frequencies between 10 kHz and 30 MHz*. ITU, 1992.
15. ITU-R Rec. P.1546-1. *Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz*. ITU, 2003.



16. **Pertsov S.V., Shchutskaya K.A.** *The input circuit of the radio*. Moscow: Energiya Publ., 1973. (rus)

17. **Golovin O.V.** *Decameter connection*. Moscow: Radio i Svyaz Publ., 1990. 240 p. (rus)

18. **Borodulin R.Yu., Borobov A.A.** Analysis technique for decoupling wire antennas by the method of moments using the fulfillment of charge continuity conditions. *Informacy and Space*, 2019, No. 3, Pp. 18–24. (rus)

Received 11.09.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

БОРОБОВ Антон Андреевич

BOROBov Anton A.

E-mail: antonborobov@yandex.ru

БОРОДУЛИН Роман Юрьевич

BORODULIN Roman Yu.

E-mail: borodulroman@yandex.ru

DOI: 10.18721/JCSTCS.12407
УДК 621.316.543.1

СИНТЕЗ И РЕАЛИЗАЦИЯ МОНОЛИТНЫХ ИНТЕГРАЛЬНЫХ СХЕМ СВЧ-ПЕРЕКЛЮЧАТЕЛЕЙ НА ОСНОВЕ GaAs pHEMT-ТЕХНОЛОГИИ

А.Ф. Березняк¹, А.С. Коротков²

¹ Всероссийский научно-исследовательский институт радиоаппаратуры,
Санкт-Петербург, Российская Федерация;

² Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Предложен новый метод синтеза проектирования твердотельных СВЧ-переключателей. Метод базируется на процедуре синтеза по ФНЧ-прототипу. Синтез СВЧ-переключателя осуществляется по величине требуемой развязки. Минимальные вносимые потери достигаются с помощью последовательной согласующей индуктивности, создаваемой на поверхности кристалла интегральной схемы. Путем последовательных частотных и схемных преобразований формируется топология СВЧ-переключателя. Эффективность предлагаемого метода подтверждается изготовлением GaAs pHEMT SPDT СВЧ-монокристаллической интегральной схемы переключателя С-диапазона.

Ключевые слова: синтез, реализация, СВЧ-переключатель, AlGaAs, GaAs pHEMT, SPDT.

Ссылка при цитировании: Березняк А.Ф., Коротков А.С. Синтез и реализация монокристаллических интегральных схем СВЧ-переключателей на основе GaAs pHEMT-технологии // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 84–96. DOI: 10.18721/JCSTCS.12407

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

DESIGN METHOD AND MANUFACTURING MONOLITHIC MICROWAVE INTEGRATED CIRCUIT SWITCHES ON GaAs pHEMT

A.F. Berezniak¹, A.S. Korotkov²

¹ All-Russian Scientific Research Institute of Radio Equipment,
St. Petersburg, Russian Federation;

² Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

This paper proposes a new synthesis method for design of solid-state microwave switches. The synthesis method based on the design of a prototype low-pass filter. By implementing successive frequency transformations and circuit conversions, a switch topology obtained. The effectiveness of the method is demonstrated by manufacturing a monolithic GaAs pHEMT SPDT switch of C-band.

Keywords: design method, microwave switch, A3B5, GaAs pHEMT, SPDT.

Citation: Berezniak A.F., Korotkov A.S. Design method and manufacturing monolithic microwave integrated circuit switches on GaAs pHEMT. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 84–96. DOI: 10.18721/JCSTCS.12407

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

СВЧ-переключатели используются для коммутации СВЧ-трактов (например, антенных, приёмных, передающих, контрольно-измерительных трактов). СВЧ-переключатели входят в состав аппаратуры систем связи, радиолокации, навигации, управления [1–4]. В настоящее время развитие СВЧ-переключателей происходит по трём направлениям: рабочие частоты достигают десятков гигагерц [5–9], осваиваются новые технологии [10–12], СВЧ-переключатели входят в состав более сложных СВЧ-монолитных интегральных схем (СВЧ МИС) [13–15].

Основы общей теории СВЧ-переключателей изложены в монографиях [16, с. 2; 17, с. 19]. Существующие методы синтеза (проектирования) СВЧ-переключателей основываются на представлении переключающих элементов в виде дискретных компонентов с различными паразитными (индуктивными/ёмкостными) связями. Методы синтеза монолитных интегральных схем (МИС) переключателей, основывающиеся на представлении переключающих элементов в виде интегральных компонентов с задаваемыми после решения задачи параметрического синтеза размерами, не развиты или отсутствуют, в том числе, для МИС СВЧ-переключателей на основе полупроводников группы А3В5.

В данной статье развивается методика синтеза твердотельных СВЧ-переключателей на базе представления СВЧ-переключателя как параметрического частотного фильтра [18]. Приводятся результаты синтеза и реализации монолитной интегральной схемы СВЧ-переключателя С-диапазона типа SPDT – «один вход – два выхода» на основе отечественной GaAs pHEMT-техно-

логии. Полученные результаты подтверждают справедливость предлагаемого подхода к синтезу твердотельных СВЧ-переключателей на полевых транзисторах.

Выбор аппроксимации Баттерворта для синтеза СВЧ-переключателей

Основу предлагаемой методики синтеза СВЧ-переключателей составляет компактная модель полевого транзистора (ПТ) в режиме ключа, представляющая транзистор пассивными RLC элементами [18]. На базе данной модели транзистора создадим модель (схему) Г-образного СВЧ-переключателя. Назовем такой переключатель элементарным (рис. 1 а), поскольку данная схема состоит из двух ПТ в комплементарных состояниях.

Для элементарного СВЧ-переключателя введём два состояния: ВКЛ – транзистор Т1 находится в замкнутом состоянии, транзистор Т2 находится в разомкнутом состоянии; ВЫКЛ – транзистор Т1 находится в разомкнутом состоянии, а транзистор Т2 находится в замкнутом состоянии (см. рис. 1 а).

При синтезе СВЧ-переключателей поведение амплитудно-частотной характеристики (АЧХ) в переходной области (от полосы пропускания к полосе задержания) имеет второстепенное значение, первостепенное значение имеет монотонность АЧХ и линейность фазовой характеристики. Максимально гладкую АЧХ обеспечивает фильтр Баттерворта, наиболее линейную ФЧХ обеспечивает фильтр Бесселя. Сравнение передаточных функций второго порядка Баттерворта:

$$H(p) = \frac{1}{p^2 + \sqrt{2}p + 1}, \quad (1)$$

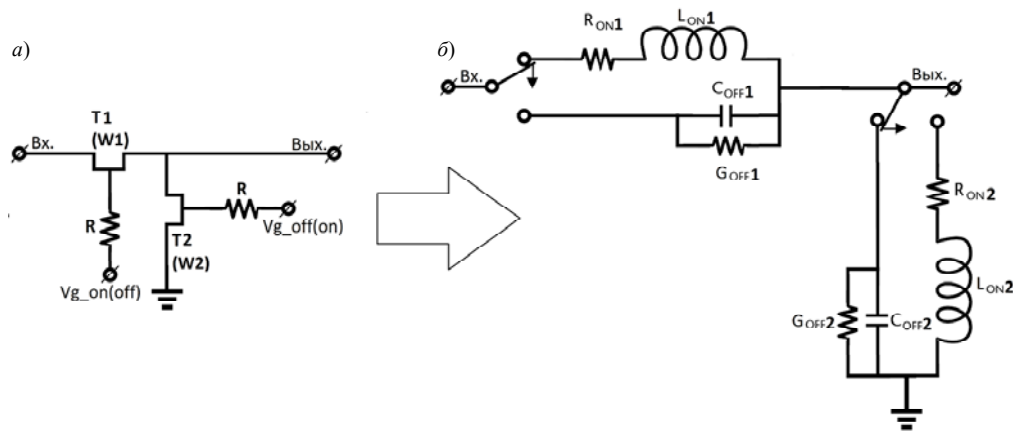


Рис. 1. Принципиальная схема Г-образного СВЧ-переключателя на ПТ – а; представление Г-образного СВЧ-переключателя через основные параметры элементов компактной модели ПТ (без учета элементов управления R) – б

Fig. 1. Schematic diagram of the Γ -shaped microwave switch on the PT – а; the representation of the Γ -shaped microwave switch through the main parameters of the elements of the compact model of the PT (excluding controls R) – б

и Бесселя

$$H(p) = \frac{1}{0,618p^2 + 1,362p + 1}, \quad (2)$$

где p – нормированная по частоте среза комплексная частота, показывает, что различия для второго порядка незначительны; коэффициенты (1 и 0,618) при p^2 ($\sqrt{2}$ и 1,362) при p близки. Поэтому в качестве аппроксимирующей функции выбираем передаточную функцию Баттерворта.

Синтез СВЧ-переключателя проведем по величине требуемой развязки, т. к. требуемые вносимые потери можно обеспечить за счет малого сопротивления «сток-исток» открытого транзистора, т. е. выбрав достаточно «широким» последовательный транзистор T1 (рис. 1 а), тогда как обеспечение требуемой развязки нуждается в нетривиальном решении.

Если в режиме ВЫКЛ амплитудно-частотную характеристику СВЧ-переключателя, представляющего в этом режиме фильтр верхних частот (ФВЧ), аппроксимировать модулем передаточной функции Баттерворта с частотой среза, равной максимальной рабочей частоте f_{p_max} , на которой модуль АЧХ равен требуемой развязке между входом и выходом СВЧ-переключателя

А_{off_min}[дБ], то получим искомое решение. Монотонность функции Баттерворта гарантирует, что в диапазоне рабочих частот ниже f_{p_max} СВЧ-переключатель обеспечит развязку выше требуемого значения, т. е. точка на АЧХ с координатами f_{p_max} , А_{off_min}[дБ] будет соответствовать минимальной развязке СВЧ-переключателя.

Подготовительный этап процедуры синтеза твердотельных СВЧ-переключателей

Для синтеза искомого ФВЧ необходимо синтезировать его ФНЧ-прототип. Воспользоваться справочными таблицами значений элементов для ФНЧ-прототипов нельзя, т. к. они нормированы по частоте среза по уровню АЧХ равному –3 дБ, тогда как требуются нестандартные уровни АЧХ (например, А_{off_min} = –20 дБ, –30 дБ и т. д.). Нагрузки на входе и выходе СВЧ-переключателя нормируем на $Z_0 = 50$ Ом. После расчета нормированного ФНЧ-прототипа Баттерворта, описанного в [18], найдем нормированную индуктивность g_1 и нормированную емкость g_2 ФНЧ-прототипа:

$$g_1 = \sqrt{2\varepsilon_p}; \quad g_2 = \sqrt{2\varepsilon_p}, \quad (3)$$

где $\varepsilon_p = \sqrt{10^{|A_{off_min}[дБ]/10|}} - 1$ – коэффициент неравномерности.

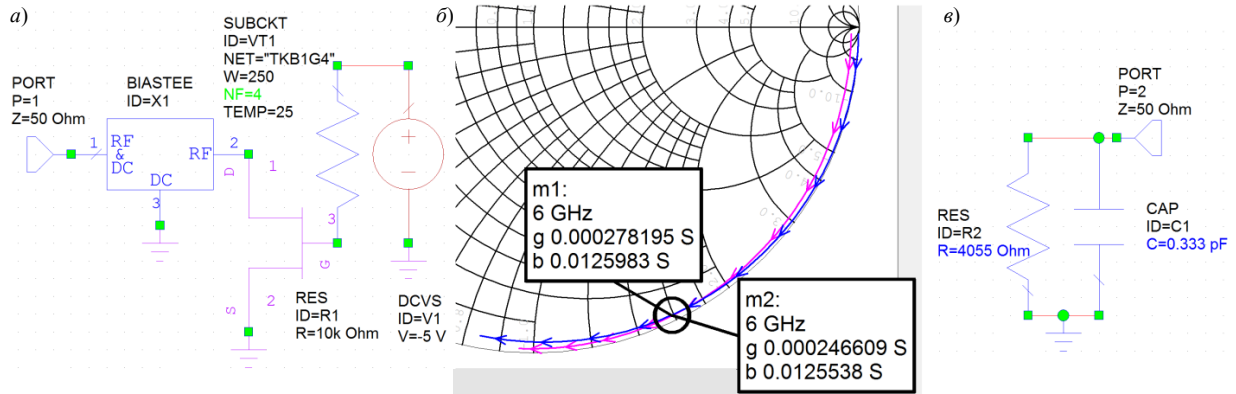


Рис. 2. Иллюстрация методики идентификации основных параметров компактной модели полевого транзистора в режиме ключа

Fig. 2. Illustration of a method for identifying the main parameters of a compact field-effect transistor model in key mode

Нормированный ФНЧ-прототип преобразуем в искомый ФВЧ с помощью преобразований вида:

$$\begin{aligned} L_{\text{ФВЧ}} &= \frac{Z_0}{2\pi f_{p_max} g_1}, \\ C_{\text{ФВЧ}} &= \frac{1}{2\pi f_{p_max} g_2 Z_0}, \\ L_{\text{ФВЧ}} &= L_{ON2} = L_{on_ud} / W_2, \\ C_{\text{ФВЧ}} &= C_{OFF1} = C_{off_ud} W_1, \end{aligned} \quad (4)$$

где $L_{\text{ФВЧ}} = L_{ON2}$ — индуктивность параллельного открытого транзистора; $C_{\text{ФВЧ}} = C_{OFF1}$ — ёмкость последовательного закрытого транзистора; W_1, W_2 — ширина затвора, последовательного и параллельного ПТ, соответственно (см. рис. 1); $L_{on_ud}[\text{нГн*мм}]$ — удельная (погонная) индуктивность «сток-исток» открытого транзистора, численно равная величине индуктивности транзистора с шириной затвора 1 мм, измеренной на рабочей частоте при последовательной схеме замещения; $C_{off_ud}[\text{пФ/мм}]$ — удельная (погонная) ёмкость «сток-исток» закрытого транзистора, численно равная величине ёмкости транзистора с шириной затвора 1 мм, измеренной на рабочей частоте при параллельной схеме замещения.

Таким образом, параметрический синтез переключателя сводится к определению значений W_1, W_2 , для чего необходимо

идентифицировать (определить) удельные величины $L_{on_ud}[\text{нГн*мм}]$ и $C_{off_ud}[\text{пФ/мм}]$.

Идентификацию проводим по S-параметрам двухполюсника (область транзистора «сток-исток» закрытого/открытого транзистора с шириной затвора равной 1 мм, рис. 2 а). S-параметры могут быть как измеренными, так и рассчитанными. В данном примере использовалась компактная модель транзистора из технологического набора конструктора (PDK) АО «Светлана-Рост» для технологии рНЕМТ05. Полученные S-параметры отображены на диаграмме Смита (рис. 2 б). Идентифицируемая модель транзистора представляет параллельное соединение RC элементов для закрытого состояния транзистора (рис. 2 в):

$$G_{off_ud} [1/(\text{Ом*мм})] \parallel C_{off_ud} [\text{пФ/мм}],$$

где $G_{off_ud}[1/\text{Ом*мм}]$ назовем удельной (погонной) проводимостью «сток-исток» закрытого транзистора, численно равной проводимости транзистора с шириной затвора 1 мм, измеренной на рабочей частоте при параллельной схеме замещения. Для открытого состояния транзистора модель представляет последовательное соединение RL элементов:

$$R_{on_ud} [\text{Ом*мм}] \text{ ser } L_{on_ud} [\text{нГн*мм}],$$

где $R_{on_ud}[\text{Ом*мм}]$ назовем удельным (погонным) сопротивлением «сток-исток» открытого транзистора, численно равным

сопротивлению транзистора с шириной затвора 1 мм, измеренному на рабочей частоте при последовательной схеме замещения.

Идентификация проводилась путем итерационного изменения величин $C1$ и $R2$ (рис. 2 в). На рис. 2 б приведен фрагмент диаграммы Смита с результатами оценки (расчета) параметра S_{11} , из которого видно, что погрешность идентификации параметров $C_{off-ud} = C1$ и $G_{off-ud} = 1/R2$ не превышает несколько процентов. Параметры L_{on_ud} и R_{on_ud} идентифицируются аналогично для транзистора в открытом состоянии.

Таким образом, в результате решения задачи идентификации параметров компактной модели полевого транзистора в режиме ключа были определены четыре параметра, которые назовем основными (табл. 1).

Для упрощения расчетов и сопоставления различных технологий с целью выбора оптимальной технологии для решения требуемой задачи синтеза введем четыре дополнительных параметра, которые назовем производными параметрами (табл. 2), поскольку они формируются из основных параметров.

Этап параметрического синтеза твердотельных СВЧ-переключателей

В режиме ВЫКЛ потерями в последовательном запертом транзисторе можно пренебречь (реактивная проводимость намного выше активной, рис. 2 б) и синтезировать последовательный транзистор (рассчитать его ширину W_1), рис. 1:

$$W_1[\text{мм}] = \frac{C_{\text{ФВЧ}}}{C_{\text{off-ud}}} = \frac{1}{2\pi f_{p_max} C_{\text{off-ud}} g_1 Z_0} = \frac{f_r}{f_r 2\pi f_{p_max} C_{\text{off-ud}} g_1 Z_0} = \frac{2\pi f_r}{\frac{1}{\sqrt{L_{on_ud} C_{\text{off-ud}}}} 2\pi f_{p_max} C_{\text{off-ud}} g_1 Z_0} = \frac{1}{K \frac{Z_0}{\rho_X} g_1} \quad (5)$$

В режиме ВЫКЛ потерями в параллельном открытом транзисторе пренебречь нельзя. Реактивное сопротивление индуктивности

$$|X_L| = 2\pi f_{p_max} L_{on_ud} = 2\pi \times 6 \times 10^9 \times 0,046 \times 10^{-9} = 1,734 \text{ Ом*мм}$$

сравнимо с активным сопротивлением открытого транзистора $R_{on_ud} = 1,575 \text{ Ом*мм}$.

Таблица 1

Основные параметры компактной модели ПТ в режиме ключа для технологии pHEMT05

Table 1

The main parameters of the compact model PT in key mode for pHEMT05 technology

Основные параметры			
R_{on_ud}	L_{on_ud}	G_{off-ud}	C_{off-ud}
1,575 Ом*мм	0,0512 нГн*мм	$2,466 \times 10^{-4} \text{ 1/Ом*мм}$	0,333 пФ/мм

Таблица 2

Производные параметры компактной модели ПТ в режиме ключа для технологии pHEMT05

Table 2

Derived parameters of the compact PT model in key mode for pHEMT05 technology

Производные параметры			
Резонансная частота	Характеристическое сопротивление	Добротность	Нормированная частота
$f_r = \frac{1}{2\pi \sqrt{L_{on_ud} C_{\text{off-ud}}}}$	$\rho_X = \sqrt{\frac{L_{on_ud}}{C_{\text{off-ud}}}}$	$Q = \frac{1}{R_{on_ud}} \sqrt{\frac{L_{on_ud}}{C_{\text{off-ud}}}}$	$K = \frac{2\pi f_{p_max}}{2\pi f_r}$
$f_r = 40,726 \text{ ГГц}$	$\rho_X = 11,771 \text{ Ом}$	$Q = 7,47$	$K = 0,1473, f_{p_max} = 6 \text{ ГГц}$

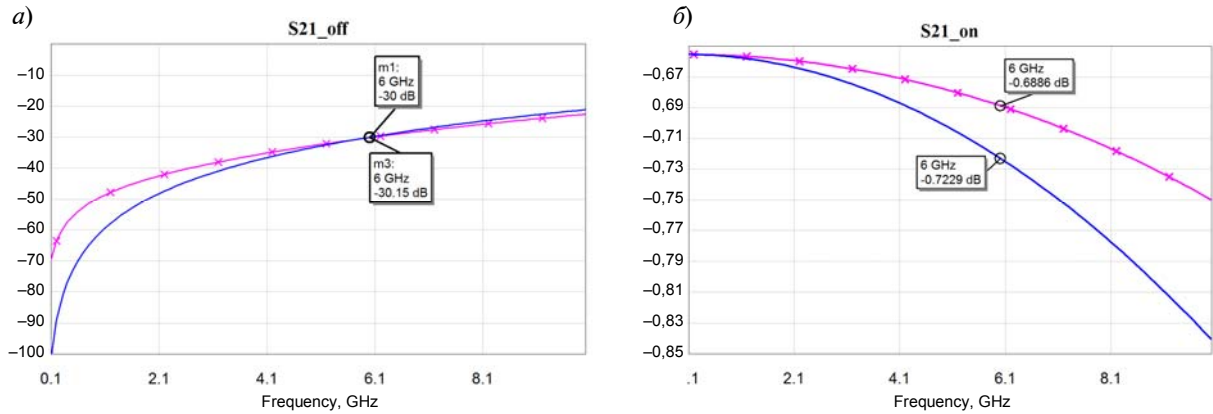


Рис. 3. АЧХ SPST СВЧ-переключателя в режиме ВЫКЛ/ВКЛ – кривые с крестиками:
а – АЧХ ФВЧ – гладкая кривая; б – АЧХ RC фильтра – гладкая кривая

Fig. 3. Frequency response of the SPST microwave switch in the OFF/ON mode – curves with crosses:
а – the frequency response of the high-pass filter is a smooth curve; б – Frequency response of the RC filter – smooth curve

Ширину параллельного транзистора находим, считая модуль реактивного сопротивления

$$|X_L| = 2\pi f_{p_max} L_{ФВЧ} \quad (6)$$

на максимальной рабочей частоте f_{p_max} равным модулю комплексного сопротивления

$$|Z_{LR}| = \frac{\sqrt{(R_{on_ud})^2 + (K\rho_x)^2}}{W_2} = \frac{R_{on_ud}\sqrt{1+K^2Q^2}}{W_2}, \quad (7)$$

$$2\pi f_{p_max} L_{ФВЧ} = \frac{R_{on_ud}\sqrt{1+K^2Q^2}}{W_2}, \quad (8)$$

$$\frac{Z_0}{g_2} = \frac{R_{on_ud}\sqrt{1+K^2Q^2}}{W_2}, \quad (9)$$

$$W_2 = \frac{g_2 R_{on_ud} \sqrt{1+K^2Q^2}}{Z_0}. \quad (10)$$

Проиллюстрируем приведенный алгоритм синтеза SPST СВЧ-переключателя численным примером. Пусть требуется синтезировать SPST СВЧ-переключатель, обеспечивающий на частоте $f_{p_max} = 6$ ГГц развязку равную или больше $A_{off_min} = -30$ дБ.

1. Найдем ФНЧ-прототип фильтра Баттерворта, т. е. для требуемой развязки

$A_{off_min} = -30$ дБ найдем нормированную индуктивность и ёмкость согласно (3):

$$g_1 = g_2 = \sqrt{2\varepsilon_p} = \sqrt{2\sqrt{10^{|A_{off_min}|[дБ]/10}} - 1} = \sqrt{2\sqrt{999}} = 7,95.$$

2. Найдем геометрические размеры последовательного ПТ (ширину затвора рНЕМТ транзистора) согласно (5) (см. табл. 2):

$$W_1 = \frac{C_{ФВЧ}}{C_{off_ud}} = \frac{1}{K \frac{Z_0}{\rho_x} g_1} = \frac{1}{0,1473 \frac{50}{11,771} 7,95} = 0,201 \text{ мм}$$

3. Найдем геометрические размеры параллельного ПТ (ширину затвора рНЕМТ транзистора) согласно (10):

$$W_2 = \frac{g_2 R_{on_ud} \sqrt{1+K^2Q^2}}{Z_0} = \frac{7,95 * 1,575 * 1,487}{50} = 0,372 \text{ мм}.$$

Верифицируем полученный результат с помощью САПР AWR MWO, т. е. построим зависимости АЧХ ФВЧ и АЧХ SPST СВЧ-переключателя в режиме ВЫКЛ при рассчитанных значениях элементов (рис. 3 а).

На рис. 3 гладкая кривая соответствует АЧХ ФВЧ, элементы которого рассчитаны согласно (4):

$$L_{ФВЧ} = \frac{Z_0}{2\pi f_{p_max} g_1} = \frac{50}{2\pi * 6 * 10^9 * 7,95} = 0,16683 \text{ нГн},$$

$$C_{ФВЧ} = \frac{1}{2\pi f_{p_max} g_2 Z_0} = \frac{1}{2\pi * 6 * 10^9 * 7,95 * 50} = 0,06673 \text{ пФ}.$$

Кривая с крестиками соответствует АЧХ СВЧ-переключателя в режиме ВЫКЛ, рассчитанной с помощью компактной модели (см. рис. 1), элементы которой определены как

$$C_{OFF1} = C_{off_ud} W_1 = 0,333 \times 0,201 = 0,066933 \text{ пФ},$$

$$G_{OFF1} = G_{off_ud} W_1 = 2,466 \times 10^{-4} \times 0,201 = 4,9567 \times 10^{-5} \text{ Ом}^{-1} = 1 / 20,175 \text{ кОм},$$

$$R_{ON2} = R_{on_ud} / W_2 = \frac{1,575}{0,372} = 4,234 \text{ Ом},$$

$$L_{ON2} = L_{on_ud} / W_2 = \frac{0,0512}{0,372} = 0,1376 \text{ нГн}.$$

На этом этап параметрического синтеза СВЧ-переключателя в режиме ВЫКЛ для элементарного СВЧ-переключателя, обеспечивающего развязку, равную 30 дБ на частоте 6 ГГц, завершен. Найдены размеры транзисторов Т1 и Т2 ($W_1 = 0,201$ мм, $W_2 = 0,372$ мм), которые при изготовлении по технологии рНЕМТ05 (табл. 1) обеспечат требуемую развязку.

Этап структурного синтеза твердотельных СВЧ-переключателей

Рассмотрим этап структурного синтеза СВЧ-переключателя в режиме ВКЛ. Необходимо оценить максимальные потери, вносимые синтезированным переключателем в режиме ВКЛ A_{on_max} [дБ] на частотах до $f_{p_max} = 6$ ГГц, для чего воспользуемся известным выражением [19, 20], в котором не учитывается влияние индуктивности

L_{ON1} транзистора в открытом состоянии (см. рис. 1):

$$A_{on_max} [\text{дБ}] = 10 \lg \left[\left(1 + \frac{R_s}{2Z_0} \right)^2 + \left(\frac{Z_0 + R_s}{2X_c} \right)^2 \right], \quad (11)$$

где $R_s = \frac{R_{on_ud}}{W_1}$ — сопротивление последовательного переключающего элемента, $X_c = \frac{1}{2\pi f_{p_max} C_{off_ud} W_2}$ — реактивное сопротивление параллельного переключающего элемента.

В рассмотренном примере переключатель, обеспечивающий развязку 30 дБ, будет вносить потери

$$R_s = \frac{R_{on_ud}}{W_1} = \frac{1,575}{0,201} = 7,836 \text{ Ом};$$

$$X_c = \frac{1}{2\pi f_{p_max} C_{off_ud} W_2} = \frac{1}{2\pi * 6,0e9 * 0,333e-12 * 0,372} = 214,09 \text{ Ом};$$

$$A_{on_max} [\text{дБ}] = 10 \lg \left[\left(1 + \frac{R_s}{2Z_0} \right)^2 + \left(\frac{Z_0 + R_s}{2X_c} \right)^2 \right] = 10 \lg [1,163 + 0,0182] = 0,723 \text{ дБ}.$$

Если учесть индуктивность открытого транзистора

$$L_{ON} = \frac{R_{on_ud}}{W_1} = \frac{0,0512}{0,372} = 0,1376 \text{ нГн},$$

получим меньшие вносимые потери, рис. 3 б), (АЧХ SPST СВЧ-переключателя в режиме ВКЛ, рассчитанная по компактной модели — кривая с крестиками; эта же АЧХ, рассчитанная по выражению (11) [19, 20], — гладкая кривая). Полученное «уменьшение» вносимых потерь указывает на возможность компенсации потерь, обусловленных отражением, путем согласования синтезированного элементарного СВЧ-переключателя по входу и выходу. К этому



же выводу придем, если исследовать зависимость вносимых потерь и развязки элементарного СВЧ-переключателя от размеров последовательного W_1 и параллельного W_2 транзисторов (табл. 3).

Таблица 3

Топологические и электрофизические параметры синтезированных СВЧ-переключателей

Table 3

Topological and electrophysical parameters of the synthesized microwave switches

$A_{off\ min}$, дБ	W_1 , мм	W_2 , мм	$A_{on\ max}$, дБ
20	0,716	0,160	0,196
30	0,402	0,286	0,343
40	0,226	0,508	0,616
50	0,127	0,904	1,121
60	0,071	1,607	2,108

Из таблицы видно, что чем больше требуется развязка, тем «уже» будет последовательный транзистор. «Узкий» последовательный транзистор приводит к большим вносимым потерям (большое R_{on}) и к увеличению коэффициента стоячей волны напряжения КСВН как на входе, так и на выходе СВЧ-переключателя. Следовательно, необходимо внести коррекцию на эта-

пе структурного синтеза, поскольку предложенная процедура синтеза Г-образного SPST СВЧ-переключателя позволяет синтезировать переключатель с заданной величиной развязки, но не с минимальными вносимыми потерями и КСВН. Для завершения процедуры синтеза необходим этап согласования [21, с. 285], на котором полученные вносимые потери можно уменьшить за счет уменьшения потерь на отражение, добавив новый структурный элемент: индуктивность L_{AD} .

Рассмотрим два варианта структурного синтеза: вариант с Т-образной L - C - L схемой согласования на выходе СВЧ-переключателя (рис. 4 а) и вариант с Г-образной (на входе и выходе) L - C - L схемой согласования в комбинации с внутренними реактивностями транзисторов (рис. 4 б). Если ёмкость C_{AD} в схеме согласования заменить закрытым транзистором, то в режиме ВЫКЛ СВЧ-переключателя получим дополнительное увеличение развязки.

Дополнительное увеличение развязки позволяет уменьшить исходное требование к величине развязки $A_{off\ min}$ [дБ] на полученную избыточную величину и повторить процедуру синтеза для скорректированной величины $A_{off\ min}$ [дБ]. На этом синтез МИС SPST-переключателя завершен.

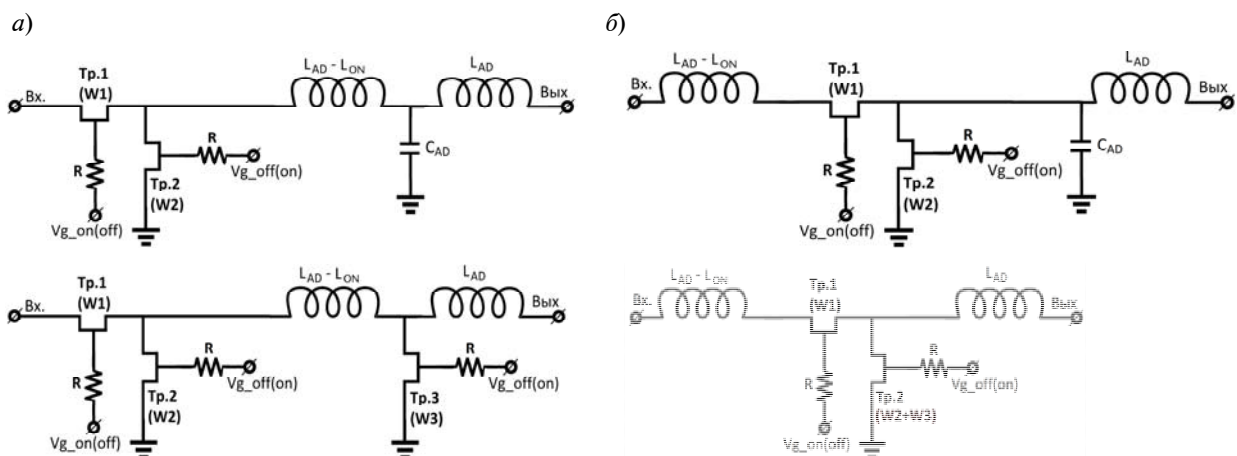


Рис. 4. Схемы согласования СВЧ-переключателя:
а – Т-образное согласование; б – Г-образное согласование

Fig. 4. Microwave switch matching circuitry

Синтез SPnT СВЧ-переключателя, где n — число выходов переключателя, проводится параллельным подключением ко входу SPST-переключателя в состоянии ВКЛ ($n - 1$) штук SPST-переключателей в состоянии ВЫКЛ. Входной импеданс SPST-переключателя в состоянии ВЫКЛ представляет ёмкость с потерями (основная часть ёмкости обеспечивается последовательным закрытым транзистором, основная часть потерь обеспечивается параллельным открытым транзистором). Реактивные потери ($n - 1$) SPST-переключателей компенсируются добавлением ко входу СВЧ-переключателя последовательной катушки индуктивности. Активные потери ($n - 1$) SPST-переключателей в состоянии ВЫКЛ приводят к увеличению вносимых потерь в n -м SPST-переключателе в состоянии ВКЛ.

По изложенному алгоритму синтеза SPnT СВЧ-переключателя была синтезиро-

вана МИС SPDT СВЧ-переключателя для С-диапазона частот. Согласование на выходе переключателя (SPST-переключателя в состоянии ВКЛ) осуществлялось с помощью Т-образной схемы согласования с транзистором 3 (рис. 4 а). Согласование на входе переключателя осуществлялось с помощью последовательной катушки индуктивности. На рис. 5 приведены параметры разработанной схемы.

МИС СВЧ-переключателя С-диапазона частот

Для экспериментального подтверждения предложенной процедуры синтеза спроектирована и изготовлена по технологии DpHEMT05 АО «Светлана-Рост» МИС SPDT СВЧ-переключателя. Данный переключатель спроектирован так, чтобы вносимые потери составляли не более 0,8 дБ, а КСВН < 1,5, тогда как развязка достигала максимума, определяемого технологией DpHEMT05.

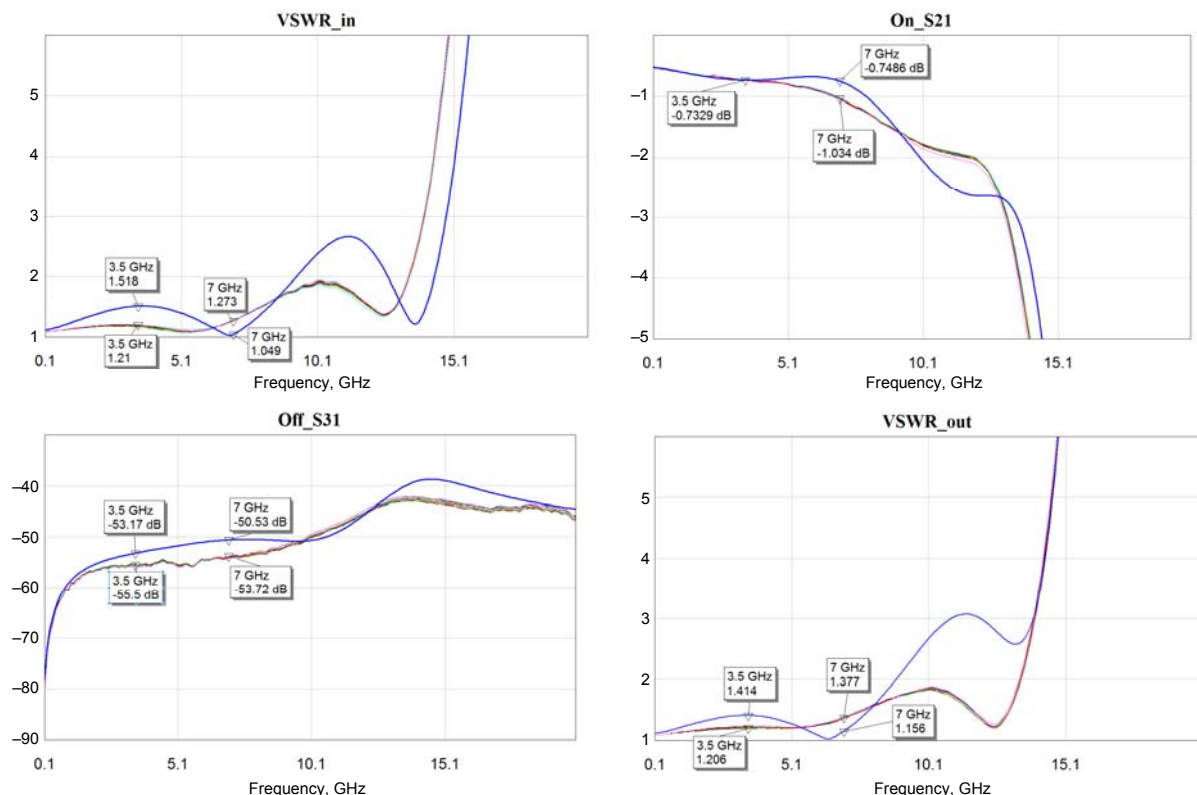


Рис. 5. Параметры МИС СВЧ SPDT-переключателя:
рассчитанные — гладкая кривая; измеренные — набор совпадающих кривых

Fig. 5. Parameters of the MIS microwave SPDT switch:
calculated — smooth curve; measured — a set of matching curves

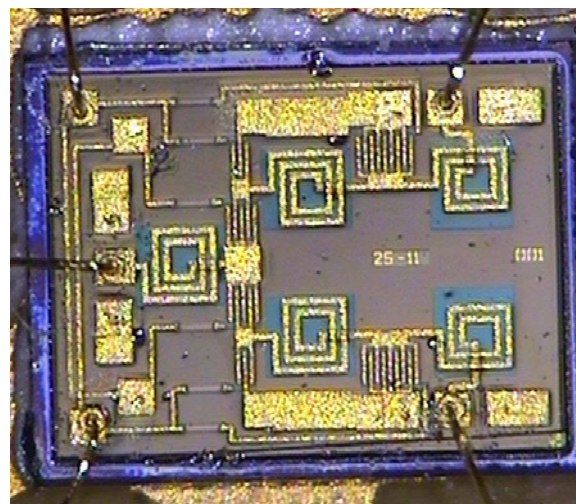
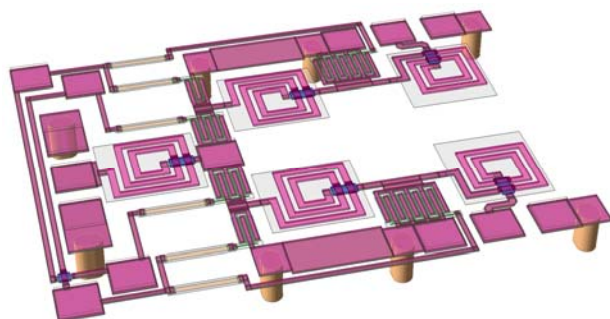


Рис. 6. Топология и фотография кристалла СВЧ МИС SPDT-переключателя С-диапазона
Fig. 6. Topology and crystal photograph of a microwave MIS SPDT C-band switch

На рис. 5 приведены параметры СВЧ-переключателя, рассчитанные (гладкая кривая) и измеренные (набор совпадающих кривых). Вносимые потери (расчетные и измеренные на графике On_S21) в открытом плече СВЧ-переключателя SPDT на частотах до 3,5 ГГц сравнимы с потерями в зарубежных аналогах (табл. 4). На более высоких частотах имеется расхождение с расчетом $> 30\%$, что связано с низкой добротностью реализованных катушек индуктивности по сравнению с добротностью, заявленной в PDK. Измеренный

КСВН на входе и выходе открытого плеча СВЧ-переключателя SPDT лучше рассчитанного КСВН благодаря низкой добротности реализованных катушек индуктивности и сравним с КСВН в зарубежных аналогах. Развязка между входом и выходом Off_S31 закрытого плеча СВЧ-переключателя SPDT (расхождение с расчетом менее 6%) превосходит развязки, приведенные для зарубежных аналогов.

На рис. 6 представлена топология кристалла и фотография МИС SPDT СВЧ-переключателя.

Таблица 4

Характеристики зарубежных аналогов изготовленного SPDT

Table 4

Characteristics of foreign analogues manufactured by SPDT

Наименование	Вносимые потери, дБ	Развязка, дБ	КСВН	Входная P_{1dB} , дБм
SKY13276-334 ¹	0,70	21	$\leq 1,2$	30 ($V_{ynp} = 3\text{ В}$) 34 ($V_{ynp} = 5\text{ В}$)
MASW6010G ²	0,8	22	$\leq 1,9$	27 ($V_{ynp} = -5\text{ В}$) 33 ($V_{ynp} = -8\text{ В}$)
RFSW8009 ³	0,65	28	$\leq 1,4$	34 ($V_{ynp} = 3\text{ В}$)
Данная статья	0,8	53	$\leq 1,2$	27 ($V_{ynp} = -5\text{ В}$) 30 ($V_{ynp} = -7\text{ В}$)

¹ Product Data Sheet SKY13276-334. Available: <https://www.rfmw.com/products/detail/sky13276334-skyworks-solutions-inc/313020/> (Accessed: 30.10.2019).

² Product Data Sheet MASW6010G. Available: <https://www.qorvo.com/products/p/RFSW8009> (Accessed: 30.10.2019).

³ Product Data Sheet RFSW8009. Available: <https://www.qorvo.com/products/p/RFSW8009> (Accessed: 30.10.2019).

Имеются сообщения о СВЧ-переключателях SPDT с развязкой, равной развязке данного изготовленного переключателя, например [22, 23], но рабочая полоса частот при этом меньше (до 4 ГГц) [22], либо вносимые потери значительно больше (1,7 дБ) [23]. Таким образом, предлагаемая методика синтеза позволяет не только формализовать процедуру расчета МИС СВЧ-переключателя, но и получить наилучшие параметры переключателя, достижимые при данной технологии.

Выводы

Предложена методика синтеза МИС СВЧ-переключателей на основе теории структурного и параметрического синтеза частотно-избирательных цепей — фильтров. При решении задачи структурного синтеза предложено расширить элементный базис схем переключателей введением дополнительного элемента — индуктивности. Это позволило свести задачу к параметрическому синтезу схемы переключателя при известном решении задачи структурного синтеза.

Разработана линейная компактная модель полевого транзистора в режиме ключа, предназначенная для синтеза СВЧ-переключателей. Параметры модели позволяют провести оценку параметров технологии изготовления, достаточных для решения задачи реализации СВЧ-переключателей с заданными линейными параметрами. Особенностью предложенной модели является возможность идентификации по S-параметрам. S-параметры могут быть получены

либо в результате измерений, либо в результате расчета по компактной модели транзистора, имеющейся в PDK фирмы-изготовителя СВЧ МИС.

Разработана методика синтеза МИС СВЧ-переключателей на полевых транзисторах, обеспечивающая заданную развязку при минимальных вносимых потерях и минимальном КСВН. Данная методика основывается на классической методике синтеза фильтра верхних частот со следующими существенными отличиями: выбор частоты среза АЧХ ФНЧ-прототипа осуществляется по максимальной рабочей частоте СВЧ-переключателя, на которой уровень затухания соответствует требуемой развязке СВЧ-переключателя; угловая частота ФНЧ-прототипа нормируется по выбранной частоте среза. Минимальные вносимые потери и минимальный КСВН обеспечиваются с помощью Г-образной или Т-образной схемы согласования на выходе МИС СВЧ-переключателя, интегрированной на кристалле с переключающими элементами.

Синтезирован и реализован СВЧ-переключатель SPDT С-диапазона по технологии GaAs pHEMT. Измеренные характеристики (вносимые потери, КСВН, развязка) близки к расчетным (расхождение с расчетом менее 6 % для развязки между входом и выходом) и сравнимы или превосходят характеристики зарубежных аналогов. При управляющем напряжении -5 В мощность точки компрессии на входе $IP_{1dB} = 27$ дБм, при управляющем напряжении -7 В — $IP_{1dB} = 30$ дБм.

СПИСОК ЛИТЕРАТУРЫ

1. Березняк А.Ф., Коротков А.С. Твердотельные СВЧ-переключатели: схемотехника, технологии изготовления, тенденции развития. Обзор. Часть 1 // Радиоэлектроника. Известия вузов. 2013. Т. 56. № 4. С. 3–28.
2. Березняк А.Ф., Коротков А.С. Твердотельные СВЧ-переключатели: схемотехника, технологии изготовления, тенденции развития. Обзор. Часть 2 // Радиоэлектроника. Известия вузов. 2013. Т. 56. № 5. С. 3–20.
3. Кочемасов В., Рауткин Ю. Интегральные СВЧ-переключатели. Часть 2 // Электроника: НТБ. 2018. № 5 (00176). С. 152–163. DOI: 10.22184/1992-4178.2018.176.5.152.163
4. Кочемасов В., Рауткин Ю. Интегральные СВЧ-переключатели. Часть 3 // Электроника: НТБ. 2018. № 6 (00177). С. 80–93. DOI: 10.22184/1992-4178.2018.177.6.80.93
5. Guo D., Qiao T., Luo X., Li M. Design of a Ka-band broadband SPDT switch MMIC based on GaN HEMTs // 2015 IEEE 16th Internat. Conf. on Communication Technology. Hangzhou, Oct. 2015. DOI: 10.1109/ICCT.2015.7399832
6. Zhao L., Liang W.F., Zhou J.Y., Jiang X. Compact 35–70 GHz SPDT switch with high isolation for high power application // IEEE Microwave and Wireless Components Letters. 2017. Vol. 27. Iss. 5. Pp. 485–487. DOI: 10.1109/LMWC.2017.2690834

7. **Trinh K.T., Kao H.L., Chiu H.C., Karmakar N.C.** A Ka-band GaAs MMIC traveling-wave switch with absorptive characteristic // *IEEE Microwave and Wireless Components Letters*. 2019. Vol. 29(6). Pp. 394–396. DOI: 10.1109/LMWC.2019.2913507
8. **Cetindogan B., Ustundag B., Turkmen E., Wietstruck M., Kaynak M., Gurbuz Y.** A D-Band SPDT switch utilizing reverse-saturated SiGe HBTs for dicke-radiometers // *11th German Microwave Conf. Freiburg, Germany, 2018*. Pp. 47–50. DOI: 10.23919/GEMIC.2018.8335025
9. **Li L., Qian R., Sun X.-W.** W-band single-pole four-throw switch for multichannel high power transceiver chipset design // *Progress in Electromagnetics Research M*. 2019. Vol. 81. Pp. 107–116.
10. **Chen C., Xu X., Yoshimasu T.** A DC–50 GHz, low insertion loss and high P1 dB SPDT switch IC in 40-nm SOI CMOS // *2017 IEEE Asia Pacific Microwave Conf. Kuala Lumpur, Malaysia, Nov. 2017*. DOI: 10.1109/APMC.2017.8251363
11. **Bhandarkar S., Nakhate S.** Asymmetric inductive substrate bias RF SPDT switch // *Materials Today Proc.* 2017. Vol. 4. Iss. 9. Pp. 10351–10355 // URL: <https://doi.org/10.1016/j.matpr.2017.06.379>
12. **Ha B.W., Seo C.W., Cho C.S., Kim Y.J.** Wideband high-isolation SPDT RF switch in 0.18 μm SiGe BiCMOS technology // *Analog Integrated Circuits and Signal Processing*. 2016. Vol. 87. Iss. 1. Pp 11–19.
13. **Hühn F., Wentzel A., Heinrich W.** Highly compact GaN-based all-digital transmitter chain including SPDT T/Rx switch for massive MIMO applications // *Internat. J. of Microwave and Wireless Technologies*. 2019. Vol. 11. Special Issue 7. Pp. 609–617 // URL: <https://doi.org/10.1017/S175907871900045X>
14. **Truong N.K., Lee D.-S., Kim S.-J., Lee M.** 40 dB-isolation, 1.85 dB-insertion loss full CMOS SPDT switch with body-floating technique and ultra-small active matching network using on-chip solenoid inductor for BLE applications // *Electronics*. 2018. Vol. 7. P. 297. DOI: 10.3390/electronics7110297
15. **Memioğlu O., Kazan O., Turan I., Karakuzulu A., Gundel A., Kocer F., Civi O.A.** Development of X-band transceiver MMIC's using GaN technology // *Advanced Electromagnetics*. 2019. Vol. 8. No. 2. Pp. 1–9 // URL: <https://doi.org/10.7716/aem.v8i2.1012>
16. **Caverly R.H.** Microwave and RF semiconductor control device modeling. Boston and London: Artech House, 2016.
17. **Bahl I.J.** Control components using Si, GaAs, and GaN technologies. Boston and London: Artech House, 2014.
18. **Березняк А.Ф., Коротков А.С., Балашов Е.В.** Трансформация синтеза фильтров частот в синтез твердотельных СВЧ-переключателей // *Наноиндустрия*. 2017. Т. 74 (Спецвыпуск). С. 427–432.
19. **Cory R., Fryklund D.** Solid state RF/Microwave switch technology: Part 2 // *MPD/Microwave Product Digest*, June 2009 // URL: https://www.rfmw.com/datasheets/skyworks/4b_solid_state_rfmicrowave_switchtechnology_part2.pdf (Дата обращения: 21.11.2019).
20. **Microsemi-Watertown.** The pin diode circuit designers handbook. Watertown: Microsemi Corp.-Watertown, 1998 // URL: https://www.ieee.li/pdf/essay/pin_diode_handbook.pdf (Дата обращения: 21.11.2019).
21. **Poole C., Darwazeh I.** Microwave active circuit analysis and design. London: Academic Press, 2016.
22. **Wang G.Q., Liu C.P.** Wideband SPDT switch with TTL control // *Applied Mechanics and Materials online since: Aug. 2014*. Vol. 599–601. Pp. 1820–1823. URL: <https://doi.org/10.4028/www.scientific.net/AMM.599-601.1820>
23. **Deng J., Gan X.** A X-band SPDT switch // *3rd Workshop on Advanced Research and Technology in Industry Applications. Advances in Engineering Research*. 2017. Vol. 148. Pp. 453–456.

Статья поступила в редакцию 25.10.2019.

REFERENCES

1. **Berezniak A.F., Korotkov A.S.** Solid-state microwave switches: circuitry, manufacturing technologies and development trends. Review (Part 1). *Radioelektronika. Izvestiya Vysshikh Uchebnykh Zavedenii*, 2013, Vol. 56, No. 4, Pp. 3–28. (rus)
2. **Berezniak A.F., Korotkov A.S.** Solid-state microwave switches: circuitry, manufacturing technologies and development trends. Review (Part 2). *Radioelektronika. Izvestiya Vysshikh Uchebnykh Zavedenii*, 2013, Vol. 56, No. 5, Pp. 3–20. (rus)
3. **Kotchmasov V., Rautkin Y.** Integrated microwave switches. Part 2. *Electronica: NTB*, 2018, No. 5 (00176), Pp. 152–163. DOI: 10.22184/1992-4178.2018.176.5.152.163 (rus)
4. **Kotchmasov V., Rautkin Y.** Integrated microwave switches. Part 3. *Electronica: NTB*, 2018, No. 6 (00177), Pp. 80–93. DOI: 10.22184/1992-4178.2018.177.6.80.93 (rus)
5. **Guo D., Qiao T., Luo X., Li M.** Design of a Ka-band broadband SPDT switch MMIC based on

- GaN HEMTs. 2015 *IEEE 16th International Conference on Communication Technology*, Hangzhou, 18–20 Oct. 2015, DOI: 10.1109/ICCT.2015.7399832
6. Zhao L., Liang W.F., Zhou J.Y., Jiang X. Compact 35–70 GHz SPDT switch with high isolation for high power application. *IEEE Microwave and Wireless Components Letters*, 2017, Vol. 27, Issue 5, Pp. 485–487. DOI: 10.1109/LMWC.2017.2690834
 7. Trinh K.T., Kao H.L., Chiu H.C., Karmakar N.C. A Ka-band GaAs MMIC traveling-wave switch with absorptive characteristic. *IEEE Microwave and Wireless Components Letter*, 2019, Vol. 29(6), Pp. 394–396. DOI: 10.1109/LMWC.2019.2913507
 8. Cetindogan B., Ustundag B., Turkmen E., Wietstruck M., Kaynak M., Gurbuz Y. A D-band SPDT switch utilizing reverse-saturated SiGe HBTs for dicke-radiometers. *2018 11th German Microwave Conference*, Freiburg, Germany, March 12–14, 2018, Pp. 47–50. DOI: 10.23919/GEMIC.2018.8335025
 9. Li L., Qian R., Sun X.-W. W-band single-pole four-throw switch for multichannel high power transceiver chipset design. *Progress in Electromagnetics Research M*, 2019, Vol. 81, Pp. 107–116.
 10. Chen C., Xu X., Yoshimasu T. A DC–50 GHz, low insertion loss and high P1dB SPDT switch IC in 40-nm SOI CMOS. *2017 IEEE Asia Pacific Microwave Conference*, Kuala Lumpur, Malaysia, 13–16 Nov. 2017. DOI: 10.1109/APMC.2017.8251363
 11. Bhandarkar S., Nakhate S. Asymmetric inductive substrate bias RF SPDT switch. *Materials Today Proceedings*, 2017, Vol. 4, Issue 9, Pp. 10351–10355. Available: <https://doi.org/10.1016/j.matpr.2017.06.379>
 12. Ha B.W., Seo C.W., Cho C.S., Kim Y.J. Wideband high-isolation SPDT RF switch in 0.18 μm SiGe BiCMOS technology. *Analog Integrated Circuits and Signal Processing*, 2016, Vol. 87, Issue 1, Pp 11–19.
 13. Hühn F., Wentzel A., Heinrich W. Highly compact GaN-based all-digital transmitter chain including SPDT T/Rx switch for massive MIMO applications. *International Journal of Microwave and Wireless Technologies*, 2019, Vol. 11, Special Issue 7, Pp. 609–617. Available: <https://doi.org/10.1017/S175907871900045X>
 14. Truong N.K., Lee D.-S., Kim S.-J., Lee M. 40 dB-isolation, 1.85 dB-insertion loss full CMOS SPDT switch with body-floating technique and ultra-small active matching network using on-chip solenoid inductor for BLE applications. *Electronics*, 2018, Vol. 7, P. 297. DOI: 10.3390/electronics7110297
 15. Memiöglu O., Kazan O., Turan I., Karakuzulu A., Gundel A., Kocer F., Civi O.A. Development of X-band transceiver MMIC's using GaN technology. *Advanced Electromagnetics*, 2019. Vol. 8, No. 2, Pp. 1–9. Available: <https://doi.org/10.7716/aem.v8i2.1012>
 16. Caverly R.H. *Microwave and RF semiconductor control device modeling*. Boston and London: Artech House, 2016.
 17. Bahl I.J. *Control components using Si, GaAs, and GaN technologies*. Boston and London: Artech House, 2014.
 18. Berezniak A.F., Korotkov A.S., Balashov E.V. Transformation of synthesis of frequency filters for synthesis of solid-state microwave switches. *Nanoindustry*, 2017, Vol. 74 (Special Issue), Pp. 427–432. (rus)
 19. Cory R., Fryklund D. Solid state RF/Microwave switch technology: Part 2. MPD/Microwave Product Digest, June 2009. Available: https://www.rfmw.com/datasheets/skyworks/4b_solid_state_rfmicrowave_switchtechnology_part2.pdf (Accessed: 21.11.2019).
 20. Microsemi-Watertown. The pin diode circuit designers handbook. Watertown: Microsemi Corp.-Watertown, 1998. Available: https://www.ieee.li/pdf/esay/pin_diode_handbook.pdf (Accessed: 21.11.2019).
 21. Poole C., Darwazeh I. *Microwave active circuit analysis and design*. London: Academic Press, 2016.
 22. Wang G.Q., Liu C.P. Wideband SPDT switch with TTL control. *Applied Mechanics and Materials* online since: August 2014, Vol. 599–601, Pp. 1820–1823. Available: <https://doi.org/10.4028/www.scientific.net/AMM.599-601.1820>
 23. Deng J., Gan X. A X-band SPDT switch. 3rd Workshop on Advanced Research and Technology in Industry Applications, *Advances in Engineering Research*, 2017, Vol.148, Pp. 453–456.

Received 25.10.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

БЕРЕЗНЯК Анатолий Федорович

BEREZNIAK Anatoliy F.

E-mail: baf1954@mail.ru

КОРОТКОВ Александр Станиславович

KOROTKOV Alexander S.

E-mail: korotkov@rphf.spbstu.ru



Моделирование вычислительных, телекоммуникационных, управляющих и социально-экономических систем

DOI: 10.18721/JCSTCS.12408
УДК 004.021

ФОРМИРОВАНИЕ ОКТОДЕРЕВА ПО ОБЛАКУ ТОЧЕК ПРИ ОГРАНИЧЕНИИ ОБЪЁМА ОПЕРАТИВНОЙ ПАМЯТИ

К.О. Беляевский

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Лазерное сканирование — современный и активно развивающийся метод дистанционного зондирования. Результат лазерного сканирования представляет собой облако точек, которое кроме координат каждой точки может содержать цвет точки и другие атрибуты. Одной из особенностей технологии лазерного сканирования, обусловившей её популярность, является возможность получения достаточно плотного облака точек, что определяет высокую точность цифрового представления геометрии объекта сканирования. В некоторых случаях облака точек могут содержать миллиарды точек, для хранения которых необходимы сотни гигабайт. Загрузка и обработка таких колоссальных объёмов данных требует больших временных и вычислительных ресурсов. Распространённым подходом является построение октодерева для ускорения операций пространственного поиска и группировки близких в пространстве точек. Использование такого октодерева совместно с внешней памятью открывает возможность ограничения объёма потребляемой оперативной памяти. В статье представлен метод построения октодерева с использованием двухуровневой системы кеширования участков облака точек. Предложены способ организации процесса построения структуры данных и способ анализа эффективности метода с помощью вычислительного эксперимента. Приведены результаты анализа эффективности предложенного метода.

Ключевые слова: трёхмерное сканирование, облако точек, структуры данных, октодерево, потребление памяти, кеширование, измерение производительности.

Ссылка при цитировании: Беляевский К.О. Формирование октодерева по облаку точек при ограничении объёма оперативной памяти // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 97–110. DOI: 10.18721/JCSTCS.12408

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

GENERATING OCTREES FROM THE POINT CLOUD IN CONDITIONS OF LIMITED RANDOM ACCESS MEMORY VOLUME

K.O. Beliaevskii

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

Laser scanning is a modern and actively developing remote sensing method. Laser scanning results are represented by point clouds, which can contain not only the coordinates of each point, but also the color of the points and other attributes. One

of laser scanning technology features, which caused its popularity, is the ability to obtain a sufficiently dense cloud of points. Thanks to this function, laser scanning allows you to determine the digital representation of the geometry of the scanned object with high accuracy. In some cases, point clouds can contain billions of points that require hundreds of gigabytes of data storage. Loading and processing such enormous amounts of data requires significant time and computational resources. A common approach is to build an octree to speed up the spatial search operations and group points close in space. Using such an octree together with external memory opens the possibility of limiting the amount of RAM consumed. This article presents the method of an octree generating using a two-tier cache system for caching sections of the point cloud. The method of organizing the process of data structure building is proposed, and the approach to the analysis of the proposed method efficiency on the basis of a computational experiment is given. The results of the proposed method efficiency analysis are presented.

Keywords: three-dimensional scanning, point cloud, data structures, octree, memory consumption, caching, performance measurement.

Citation: Beliaevskii K.O. Generating octrees from the point cloud in conditions of limited random access memory volume. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 97–110. DOI: 10.18721/JCSTCS.12408

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Анализ предметной области. Процесс обработки данных лазерного сканирования состоит из двух основных этапов: 1) загрузка данных и подготовка данных к последующей обработке; 2) обработка данных с помощью конкретного алгоритма [1]. Существующие исследования сосредоточены преимущественно на решении задач, относящихся ко второму этапу, и посвящены вопросам оптимизации обработки данных [2–4].

Часто при обработке облака точек необходимо использовать информацию о пространственных отношениях между точками, например, близость (соседство) точек или попадание точек в область определенной пространственной фигуры (куба, сферы и т. п.). В силу специфики технологии лазерного сканирования, точки в полученном облаке расположены рядом в соответствии с траекторией луча сканирования. Вместе с тем, даже зная модель сканера, предсказать распределение точек не представляется возможным, т. к. их порядок в облаке может быть изменен в результате обработки каким-либо алгоритмом или при объединении нескольких облаков точек. Пример подоб-

ных облаков приведен на рис. 1 (в верхней части рисунка точки расположены участками в результате обработки, нижняя часть рисунка иллюстрирует результат сведения данных с различных сканеров; оттенки серого цвета отражают нарастание индекса точки). Точки в облаке могут быть распределены непредсказуемым образом, и операции поиска требуют полного перебора облака точек, т. е. вычислительная сложность задачи составляет $O(n)$ [5].

Для ускорения вычислений обычно используют так называемую пространственную индексацию облака точек. Результатом подобной операции будет структура данных, описывающая объём, занимаемый облаком точек, в виде набора непересекающихся регионов, и позволяющая идентифицировать каждую точку в облаке как принадлежащую к определенному региону. Таким образом, пространственный поиск будет выполняться не во всем облаке, а в интересующем нас регионе, что позволит уменьшить сложность поиска, которая в данном случае составит $O(\log n)$ [5]. Будем называть подобную структуру данных структурой разбиения пространства.

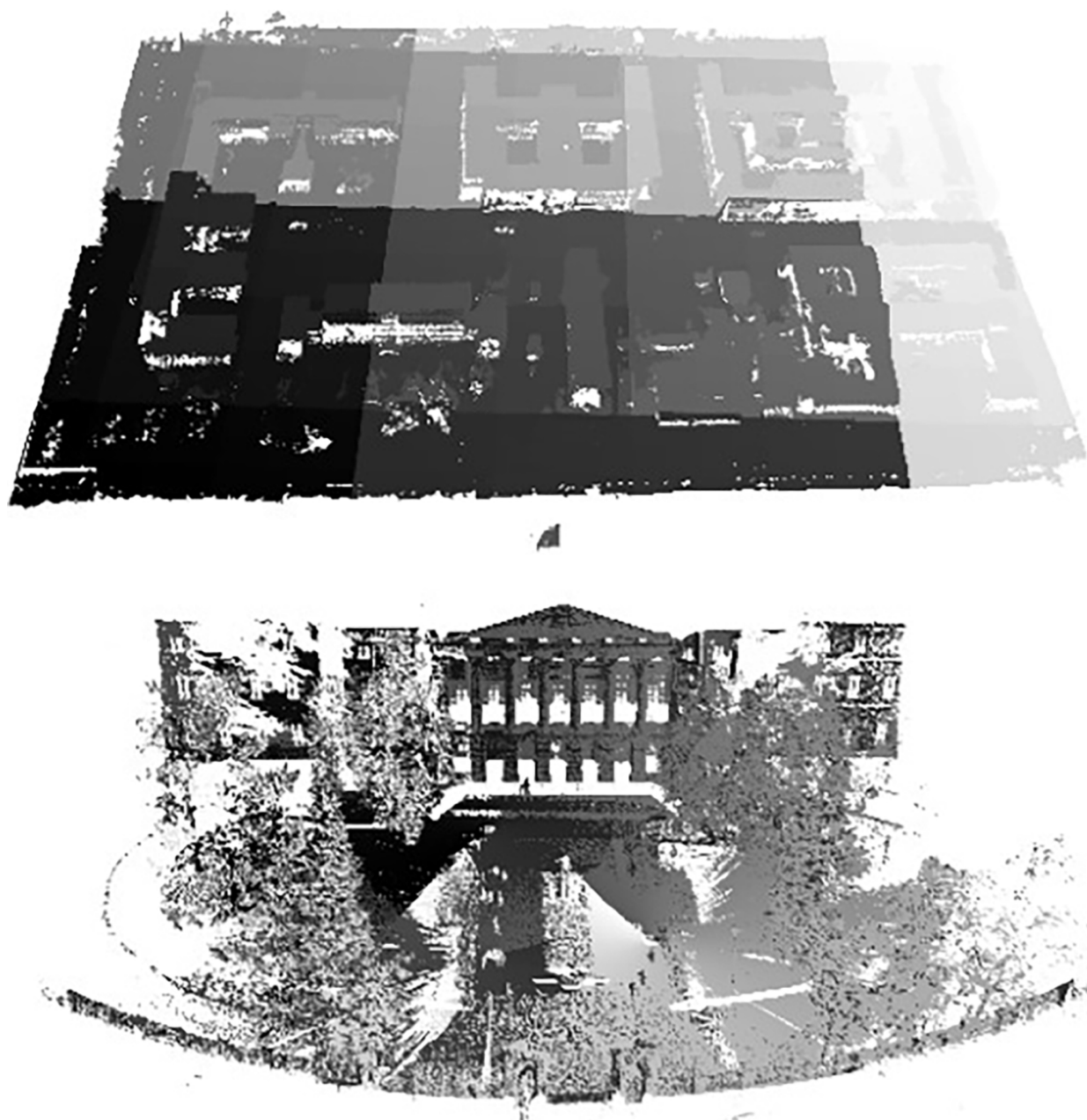


Рис. 1. Визуализация исходных индексов точек в облаке
Fig. 1. Visualization of the input indices of the point clouds

Одна из ключевых проблем при работе с облаками точек — это противоречие между большим объёмом информации, который необходимо обработать, и ограниченностью вычислительных ресурсов [6]. Размеры облаков точек нередко превышают доступные объёмы оперативной памяти вычислительных устройств, что влечет за собой необходимость обрабатывать облако точек по частям, а это, в свою очередь,

существенно усложняет процесс обработки облака точек.

Постановка задачи. В данной статье рассматривается гипотеза о том, что совместное использование структуры разбиения пространства и системы кеширования блоков позволит значительно сократить затраты оперативной памяти при обработке облака точек. В результате подобного объединения появится возможность выпол-

нять кеширование по пространственному признаку с возможностью выгрузки неиспользуемых блоков на жесткий диск.

В качестве структуры разбиения пространства в настоящей статье выбраны октатные деревья [7], ввиду наличия у них ярко выраженного разбиения облака точек на блоки. Подобное свойство позволяет производить кеширование отдельных блоков, что является более эффективным подходом, чем кеширование отдельных точек. При наличии такой системы появляется возможность загружать в оперативную память ограниченный набор точек, полностью загружать облако точек не требуется.

Октатные деревья

Октатные деревья широко используются при обработке облаков точек, полученных методом лазерного сканирования, и находят применение в таких областях, как распознавание объектов [8–10], сегментация [11, 12], выделение примитивов [13], реконструкция поверхности [14], расчет нормалей [15] и пр.

Структура октатного дерева приведена на рис. 2. Узел октатного дерева хранит в себе информацию о восьми потомках, а также вспомогательную информацию, которая может использоваться в процессе построения или обработки облака точек (напри-

мер, координаты узла, его размер, уровень и т. п.). Листья октатного дерева хранят в себе информацию о точках, которые попадают в ячейку конкретного листа на данном уровне разбиения. Листья также могут содержать вспомогательную информацию.

Формирование октатного дерева производится при помощи рекурсивного добавления точек, начиная с корневого узла. На этапе построения при добавлении точек в любой узел производится проверка на удовлетворение критерия разбиения для данного узла. В качестве критерия разбиения было выбрано превышение максимального количества точек в узле. Выбор максимального количества точек в узле обусловлен решаемой задачей, анализ возможных значений приведен в разделе «Результаты тестирования». При выполнении указанного критерия производится разбиение узла на восемь меньших октантов, точки родительского узла распределяются между ними по пространственному признаку. Критерием остановки служит превышение максимальной глубины дерева. Максимально допустимая глубина дерева составляет 21 уровень, что обусловлено способом идентификации узла при помощи 64-битного целочисленного значения (используется по 3 бита на уровень). Количество узлов на максимальном уровне разбиения составляет 8^{21} .

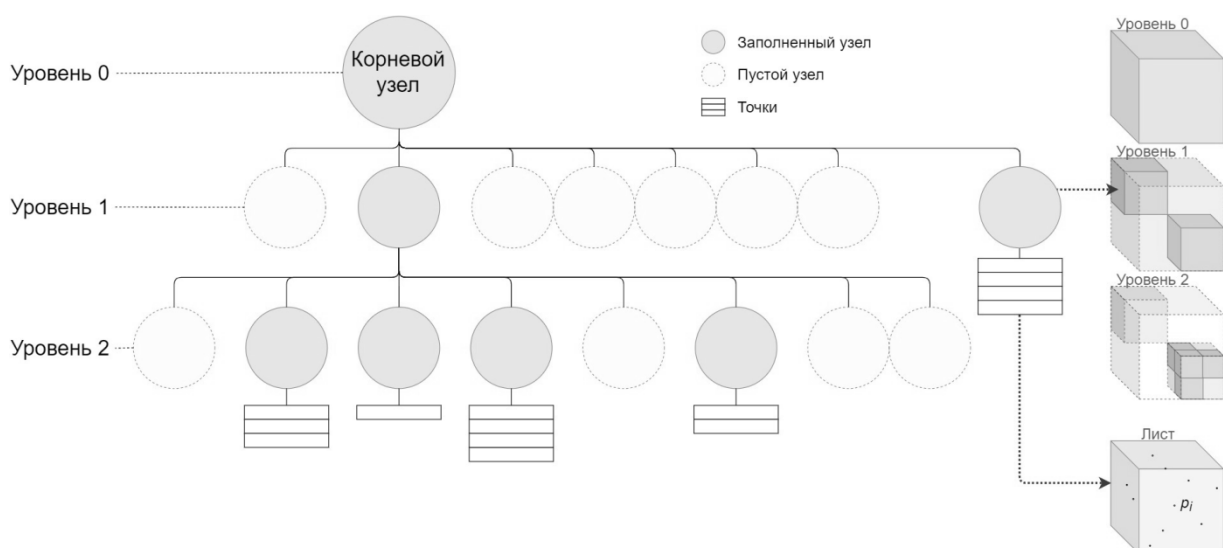


Рис. 2. Структура октатного дерева
Fig. 2. The octree structure



В большинстве опубликованных работ построение октарного дерева выполняется в оперативной памяти. Это означает, что узлы, листья и принадлежащие им точки также будут храниться в оперативной памяти. Такое решение существенно ускоряет построение дерева, однако делает невозможным обработку облаков точек, не помещающихся в оперативную память.

Использование системы кеширования в процессе формирования октарного дерева

Для разрешения указанного выше противоречия предлагается модифицировать октарное дерево для работы с блоками точек, которые могут находиться как в оперативной памяти, так и на жестком диске. Для того чтобы не усложнять реализацию октарного дерева, предлагается оригинальная система кеширования облаков точек, назначением которой будет управление размещением блоков точек. Подобная система позволит создавать/удалять, а также подгружать/выгружать блоки точек, поддерживая при этом заданный уровень потребления оперативной памяти за счет выгрузки неиспользуемых блоков.

Предлагаемая система кеширования должна обеспечивать работу с блоками точек в условиях ограничений по объёму доступной оперативной памяти. На практике это означает, что при превышении определенного лимита потребляемой памяти, часть блоков должна быть выгружена на жесткий диск. Стоит отметить, что при подобном подходе к ограничению потребляемой памяти лимитируется только та память, в которой располагаются точки из облака точек. Иерархия октодерева, а также прочие данные программы, не относящиеся к облаку точек, не подлежат ограничению.

Для реализации подобного поведения достаточно двухуровневой системы кеширования: L_1 — оперативная память; L_2 — жесткий диск.

При этом в оперативной памяти остаются только актуальные блоки, а устаревшие вытесняются на уровень L_2 , т. е. сохраняются на жестком диске. Для вытеснения блоков можно использовать различные политики: First In First Out (FIFO), Least

Recently Used (LRU), Least Frequently Used (LFU) и прочие [16]. В нашем случае хорошо подходит политика LRU, т. к. она, с одной стороны, отличается простотой реализации (обычно представлена двусвязным списком), а, с другой стороны, удовлетворяет требованиям по удалению устаревших блоков. Стоит отметить, что эффективность выбранной политики вытеснения может варьироваться в зависимости от вида операции, производимой над облаком точек.

Для обеспечения взаимодействия с октодеревом система кеширования должна поддерживать следующие операции: создание блока точек; добавление точек к блоку точек; выгрузка блока точек; удаление блока точек; финализация блока точек.

Для ускорения построения октодерева была организована параллельная обработка данных. Так, например, файловые операции, необходимые для работы кеша на уровне L_2 , выполняются в отдельном потоке. Управляющие команды поступают при помощи асинхронной очереди сообщений. Ниже приведены описания таких операций.

1. Сохранение блока. При выгрузке блока из L_1 в очередь поступает команда на сохранение данного блока в соответствующий файл, при этом основной поток не тратит время на ожидание завершения операции.

2. Загрузка блока. В случае загрузки в очередь поступает команда на загрузку файла, после чего выполняется ожидание завершения выполнения команды.

3. Финализация блока. В процессе активной работы с блоками точек они хранятся в виде отдельных файлов (в связи с возможностью изменения их размеров). По завершении работы с блоком он присоединяется к общему файлу (данная операция была названа *финализацией*). Большое количество операций с файлами небольшого размера существенно снижает производительность [17] системы, поэтому подобную реализацию в общем случае не рекомендуется использовать. В то же время в данной работе используется алгоритм построения октарного дерева, позволяющий в большинстве случаев сразу присоединить блок к финальному файлу.

Процесс построения октарного дерева также реализуется при помощи системы кеширования: при добавлении точек в узел происходит обращение к соответствующей ячейке кеша с целью добавления новых точек. Стоит отметить отличия предложенного алгоритма накопления точек в узле от классической реализации октодерева [18]: точки после разбиения не спускаются ниже по иерархии дерева, а остаются в узле. Последующие точки пропускаются через такой узел насквозь. Такой подход позволяет существенно снизить количество файловых операций за счет двух факторов: во-первых, точно известно, что записей в данный блок не будет, и можно присоединить его к финальному файлу; во-вторых, в процессе построения практически отсутствуют операции чтения и чтения/перезаписи данных (т. к. больше нет необходимости при разбиении выполнять для накопленного блока точек цепочку операций считывание-разбиение-запись). Иначе говоря, при подобном подходе разрешается хранение блоков точек в ветвях дерева, что положительно влияет на скорость построения, однако может снизить производительность операций поиска.

Пример алгоритма построения октодерева на следующей странице в листинге.

Предложенный алгоритм имеет следующие преимущества:

- отсутствие необходимости подгружать весь блок данных в L_1 при добавлении новых точек;
- возможность финализации завершённых блоков;
- возможность асинхронной работы с файлами.

Таким образом, при построении октодерева данные подгружаются только в случае финализации узла.

Методика анализа характеристик предложенного способа формирования октодерева

Для проведения анализа эффективности предложенного способа построения октодерева разработан набор тестов, представ-

ленный в табл. 1. Программная реализация тестов и октодерева выполнена на языке C++, т. к. этот язык позволяет производить низкоуровневые операции с памятью, что необходимо при работе с большими объёмами данных.

Выбор тестов обусловлен необходимостью подтверждения работы в условиях имеющихся ограничений по объёму оперативной памяти. Для анализа работоспособности системы на различных облаках точек, выбора оптимального размера узла, а также для сравнения режимов работы с кешированием и без разработан тест 1. Для доказательства ограничения потребляемой памяти, а также для демонстрации процесса формирования октарного дерева разработаны тесты 2–3. Для анализа производительности поиска по октодереву разработан тест 4.

Существует множество форматов для хранения облаков точек, например: OBJ, PLY, XYZ, PCD, LAS и многие др. [19]. В качестве входного формата облаков точек для тестирования среди них был выбран формат LAS, т. к. он широко используется и имеет открытые исходные коды [20].

Для тестирования предложенной системы было выбрано четыре облака точек. Облако точек под названием five.las выбрано для анализа поведения системы при малом количестве точек. Облако точек под названием polytech.las выбрано для исследования возможностей системы на больших облаках точек с низкой плотностью. Облако точек под названием smolny.las выбрано для исследования возможностей системы на больших облаках точек с высокой плотностью и совмещением данных с нескольких сканеров. Облако точек под названием molodezhnoe.las представляет собой наиболее сложный случай с точки зрения обработки, т. к. оно обладает сверхбольшим размером (60 Гб), а также имеет высокую плотность за счет совмещения данных сканирования. Характеристики облаков приведены в табл. 2, внешний вид облаков точек представлен на рис. 3.

```
// Пример добавления новой точки в дерево (оригинальный алгоритм оперирует блоками точек)
void InsertPoint(Node* node, Point point) {
    if(node->isFull()) // Если нельзя добавить больше точек
    {
        // Получение индекса подузла, в который попадает точка
        int childIndex = GetChildIndex(node, point);
        // Добавление точки в подузел
        InsertPoint(GetChild(node, childIndex), point);
    }
    else
    {
        // Добавление точки в узел
        AppendPoint(node, point);
        // Если узел заполнен - финализация
        if(node->isFull())
            FinalizeNode(node);
    }
}

void AppendPoint(Node* node, Point point) {
    if(*Has L1 Cache Entry*) // Если узел загружен в RAM
    {
        GetL1Entry(node)->push_back(point); // Добавление новой точки в узел
    }
    else
    {
        // Если узел не загружен в RAM - просто дописываем точку в конец файла
        // Для отправки команды используем асинхронную очередь
        // Запись будет производиться в другом потоке
        L2Queue->newCommand(AppendPointCmd(node, point));
    }
}

void FinalizeNode(Node* node) {
    // Загрузка данных в L1 и отправка на финализацию
    L2Queue->newCommand(FinalizeNodeCmd(GetL1Entry(node)));
}
}
```

Формирование октодерева с использованием системы кеширования
An octree formation with the usage of the caching system

Таблица 1

Список тестов

Table 1

The list of the tests

Название теста	Описание теста
Тест 1. Время построения	Анализ зависимости времени построения от размера узла и объёма кеша
Тест 2. Потребление памяти	Анализ потребления памяти (RAM, HDD) в процессе построения
Тест 3. Файловые операции	Анализ операций чтения/записи в процессе построения
Тест 4. Скорость поиска	Анализ зависимости скорости поиска точек в дереве от размера узла и объёма кеша

Таблица 2

Облака точек для тестирования

Table 2

Point clouds for testing

Название	Кол-во точек, млн	Размер, Гб	Плотность, точек/м ²
five.las	1,2	0,03	18641
polytech.las	51,6	1,8	267
smolny.las	126	4,3	23656
molodezhnoe.las	1560	53,4	55097

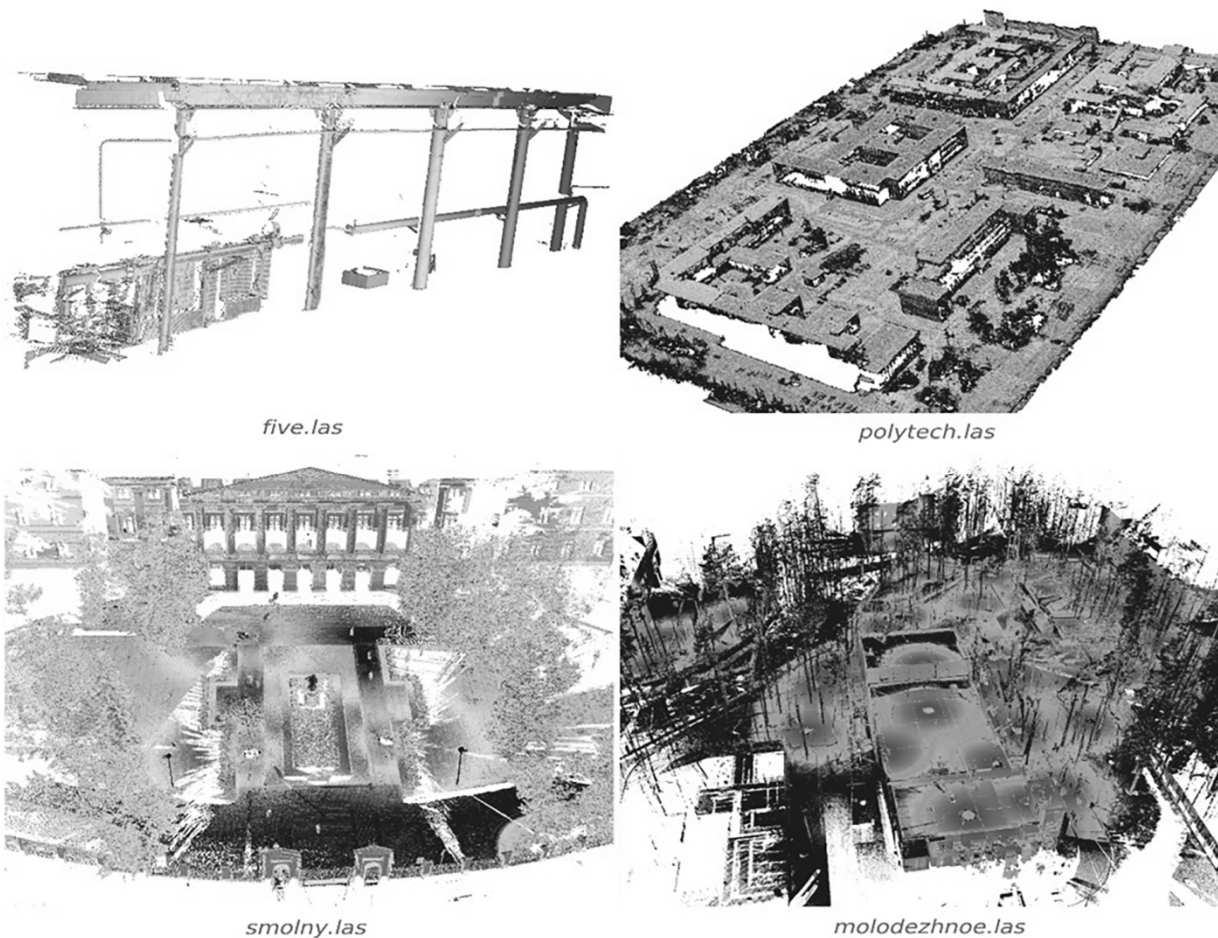


Рис. 3. Внешний вид используемых облаков точек с раскраской по плотности (шкала плотности индивидуальна для каждого облака)

Fig. 3. Visual representation of the used point clouds with density coloring (density scale is individual for each cloud)

Конфигурация тестового стенда:
 CPU: Intel® Core™ i7-7820HQ CPU @ 2.90 GHz
 RAM: 2 x SODIMM DDR4 2400 MHz (2x8 GB)
 SSD: Samsung SSD 850 PRO 1 TB

OS: Ubuntu 18.04
 File system: Ext4

Результаты тестирования

Тест 1. Время построения дерева. На рис. 4 представлены результаты анализа



производительности метода построения октодерева при использовании системы кеширования. Для этого проведены измерения времени построения дерева для различных облаков точек в оперативной (RAM) и файловой (HDD) памяти путем изменения размера узла и размера кеша. Такой тест позволит выяснить оптимальные параметры дерева, а также продемонстрировать разницу в скорости построения дерева в оперативной памяти и с использованием жесткого диска.

На графике зависимости скорости построения дерева от максимального размера узла видно, что построение дерева значительно замедляется при использовании узлов маленького размера. Это вызвано прямой зависимостью между размером узла и глубиной дерева: чем меньше узел, тем больше глубина, на которую спускается очередная точка из облака при построении. В случае с файловыми операциями размер узла влияет особенно сильно, т. к. приводит к значительному увеличению количества создаваемых файлов.

Как следует из графика зависимости скорости построения дерева от используемого объема кеша, увеличение объема кеша немного ускоряет процесс построения дерева. Тем не менее, значительного вклада это не вносит, т. к. добавляемые точки обычно разбросаны по облаку неравномерно, что понижает эффективность кеширования.

По результатам проведенного теста можно утверждать, что увеличение размера узла или объема кеша приводит к ускорению построения октодерева. Также можно заметить, что характер изменений практически не зависит от исследуемого облака точек. Однако для выбора наилучшего значения размера узла необходимо также провести измерения скорости поиска по построенному дереву. Что касается выбора подходящего размера кеша, то рекомендуется использовать максимально допустимое для конкретной системы значение, но не превышающее размер наибольшего обрабатываемого облака точек.

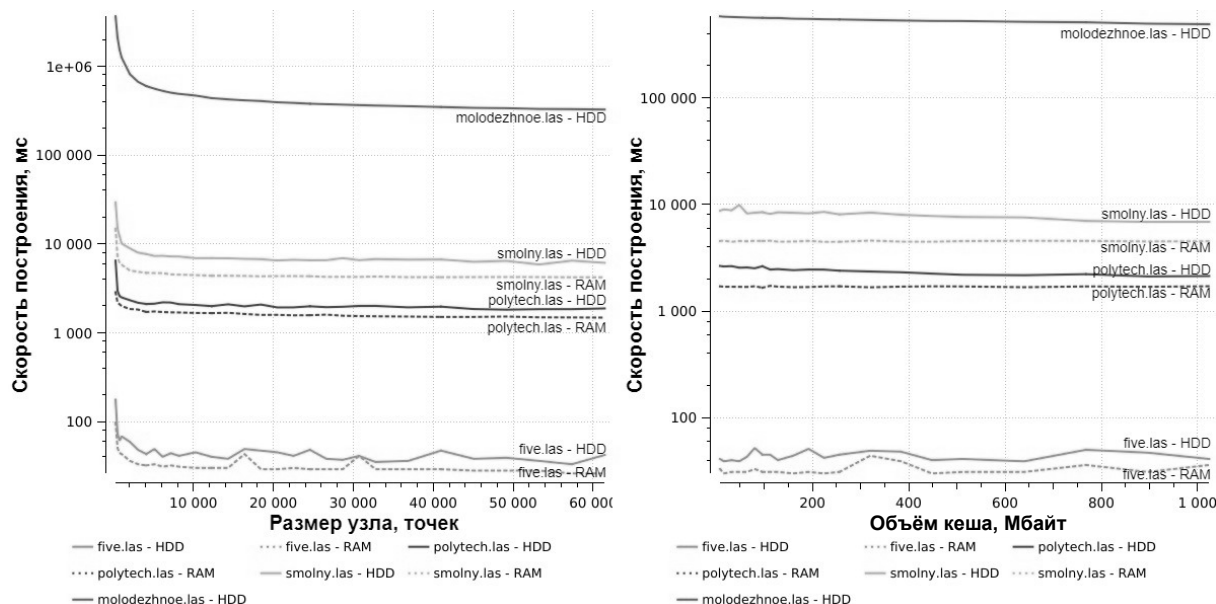


Рис. 4. Зависимость скорости построения дерева от максимального размера узла (слева) и от используемого объема кеша (справа)

Fig. 4. Dependence of the speed of an octree formation on the maximum node size (left) and on the cache volume used (right)

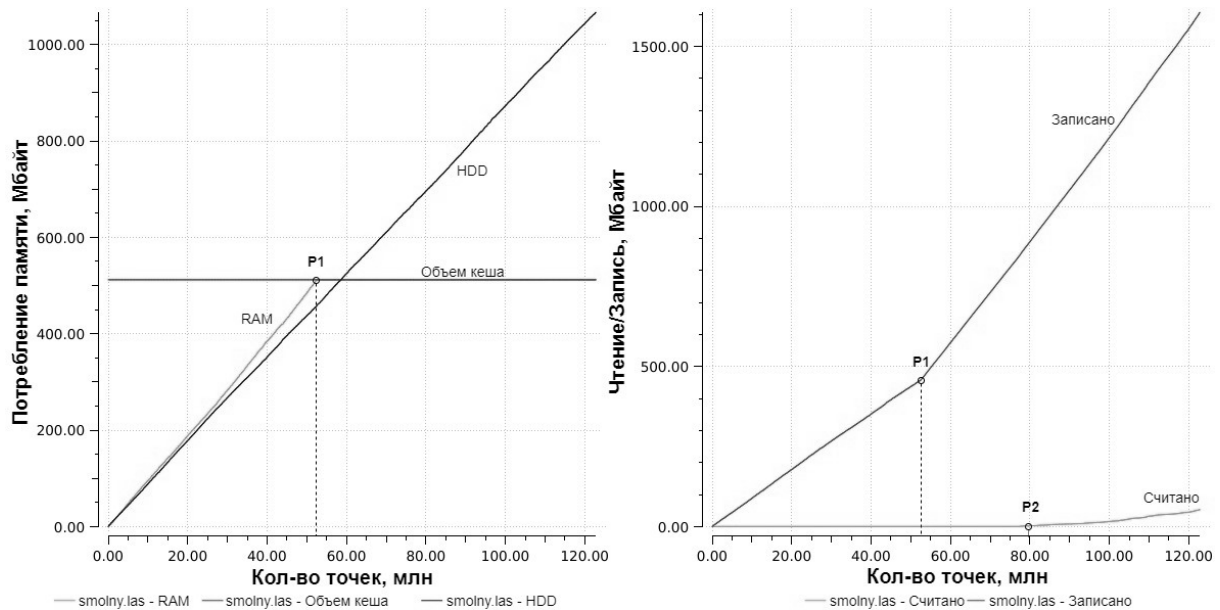


Рис. 5. Потребление памяти в процессе построения дерева (слева) и файловые операции в процессе построения дерева (справа)

Fig. 5. Memory consumption in the process of an octree formation (left) and file operations in the process of an octree formation (right)

Тесты 2 и 3. Потребление памяти и файловые операции. Далее проверялась корректность работы системы кеширования в условиях ограничения потребляемой оперативной памяти. Результаты выполнения данного теста иллюстрируются графиками потребления оперативной и файловой памяти. На рис. 5 приведены графики, снятые в процессе построения октодерева, и отражающие зависимость объёма потребления L_1 (RAM) и L_2 (HDD) от количества загруженных точек (слева) и файловые операции в процессе построения октодерева (справа).

Как видно из приведенных графиков, сначала происходит заполнение кеша до значения его ёмкости (точка P1), параллельно с выгрузкой заполненных узлов на жесткий диск. Расход оперативной памяти остается в пределах заданных значений за счет использования дискового пространства. Также на рисунке можно увидеть, как растет интенсивность записи на жесткий диск после заполнения кеша (точка P1), и как возрастает количество операций считывания по мере приближения к концу процесса построения (точка P2). Послед-

нее вызвано загрузкой незаполненных узлов, выгруженных из кеша ранее.

Результаты данного теста позволяют сделать вывод об успешности объединения октодерева и системы кеширования, т. к. потребление памяти оставалось в пределах заданных лимитов на протяжении всего процесса обработки облака точек.

Тест 4. Скорость поиска. Если в предыдущих тестах оценивались влияние размеров узла и объёма кеша на скорость построения, то в данном тесте целью будет измерение скорости поиска в уже построенном дереве. Данный тест позволит оценить, насколько замедляется скорость поиска при больших размерах узлов, а также насколько увеличение объёмов кеша ускоряет поиск.

На рис. 6 приведен график зависимости скорости поиска от размеров узлов (слева) и график зависимости скорости поиска от объёмов кеша (справа). В процессе тестирования для каждого исследуемого параметра производится большое количество операций поиска точек, находящихся в определенном радиусе от случайной точки из облака. В качестве результата используется среднее значение времени поиска.

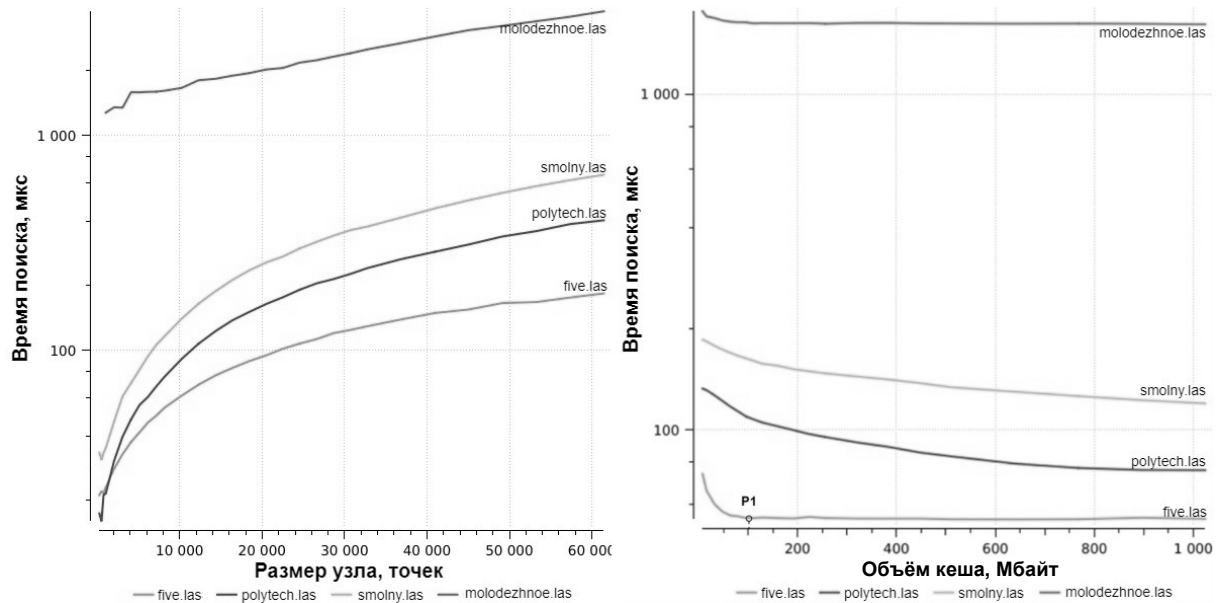


Рис. 6. Скорость поиска в зависимости от размера узла (слева) и в зависимости от объёма кеша (справа)
Fig. 6. Search speed depending on the node size (left) and depending on the cache volume (right)

Как видно из графика зависимости скорости поиска от размера узла, чем больше количество точек в узле, тем медленнее происходит поиск. Данный результат обусловлен механизмом поиска: после проверки попадания узла в радиус поиска алгоритм выполняет проверку для всех точек, входящих в узел (т. к. попадание узла не гарантирует попадание всех точек узла). Соответственно, чем больше точек в узле, тем большее количество проверок отдельных точек предстоит провести. В случае, когда размер узла невелик, большая часть точек отсекается на этапе более быстрой проверки узлов.

На графике зависимости скорости поиска от объёма кеша можно наблюдать ускорение операций поиска при увеличении объёма кеша. Стоит отметить, что оптимальный объём кеша возрастает с увеличением размера облака точек: это можно увидеть по тому, как насыщается график *five.las* в точке P1, а также по динамике изменения графиков *polytech.las* и *smolny.las*.

По результатам проведенного теста можно утверждать, что увеличение объёма кеша положительно влияет на скорость поиска, в то время как при увеличении объёма узла скорость поиска замедляется. Таким образом, при выборе подходящего размера узла следует также руководствоваться необходимой скоростью поиска. Нами было выбрано

значение 2048 точек, как компромисс между скоростью построения и скоростью поиска. Способ выбора подходящего размера кеша остается неизменным.

Сравнение с распространенными октодеревьями. Сравнение исследуемой системы с распространенными библиотеками и программами, предназначенными для обработки и визуализации облаков точек, доказало, что предложенная система не уступает по производительности. Для сравнения были выбраны следующие программные компоненты:

Point Cloud Library (PCL) [21] — библиотека с открытым исходным кодом, широко используемая для обработки облаков точек;

CloudCompare [22] — программное обеспечение с открытым исходным кодом, предназначенное для загрузки, обработки и визуализации облаков точек.

Сравнение производилось путем измерения времени, необходимого на загрузку облака точек и построения октарного дерева. Для загрузки использовались облака точек, приведенные в табл. 3. В качестве параметров нашего дерева использовались: размер узла — 2048 точек, объём кеша — 512 Мбайт. Данные параметры выбраны, как наиболее подходящие с точки зрения компромисса между скоростью построения и поиска. Результаты сравнения приведены в табл. 3.

Таблица 3

Сравнение предложенной реализации с аналогами

Table 3

Comparison of the proposed implementation with analogues

Название облака	Размер облака	Предложенная реализация		PCL		Cloud Compare	
		CPU	RAM	CPU	RAM	CPU	RAM
five.las	33 Мбайт	60 мс	33,4 Мбайт	122 мс	66 Мбайт	435 мс	114 Мбайт
polytech.las	1,8 Гбайт	2,6 с	517 Мбайт	4,6 с	1,8 Гбайт	18,2 с	1,6 Гбайт
smolny.las	4,3 Гбайт	11,6 с	586 Мбайт	12,5 с	4,3 Гбайт	58,5 с	3,6 Гбайт
molodezhnoe.las	53,4 Гбайт	14,9 мин	986 Мбайт	—	—	—	—

В результате сравнения выявлено, что предложенная оригинальная реализация октодерава значительно превосходит по скорости построения дерева остальные программные компоненты. Помимо существующих методов оптимизации в предложенной системе, подобные различия могут быть вызваны различиями в размерах узлов, политиках вставки, разрешениях дерева и прочих скрытых параметрах. При этом по времени построения дерева данная реализация, при условии кеширования на жесткий диск, как минимум не уступает приведенным аналогам, работающим в оперативной памяти.

Стоит также отметить потребление оперативной памяти: несмотря на то, что общее потребление памяти выше установленного лимита (в основном за счет данных иерархии октодерава, содержащих информацию о миллионах узлов, и не учитывающихся при вытеснении), общее потребление памяти значительно сократилось по сравнению с аналогами.

Заключение

Рассмотрена гипотеза о том, что совместное использование структуры разбиения пространства и системы кеширования блоков позволит значительно сократить затраты оперативной памяти при обработке облака точек.

Для подтверждения гипотезы представлен метод построения октодерава с использованием двухуровневой системы кеширования участков облака точек в условиях ограничений на использование

оперативной памяти, способ организации процесса построения октодерава, способ анализа эффективности метода вычислительным экспериментом.

Преимущества предложенного способа формирования октодерава показаны с использованием сравнительного анализа зависимости скорости построения и поиска по дереву от используемого размера узла дерева и объема кеша. Сравнение с существующими программными решениями показало, что предложенный способ формирования октодерава позволяет значительно сократить объем потребляемой оперативной памяти, а также не уступает по производительности построения деревьев аналогам, работающим в оперативной памяти, что подтверждает выдвинутую гипотезу.

Экспериментальное исследование позволило выявить следующие зависимости: увеличение размеров узла положительно влияет на скорость построения дерева; снижение размеров узла положительно влияет на скорость поиска по дереву; увеличение размеров кеша положительно влияет как на скорость построения, так и на скорость поиска по дереву.

Полученные результаты могут использоваться в алгоритмах обработки больших облаков точек при нехватке оперативной памяти. Стоит отметить, что ввиду зависимости скорости построения и скорости поиска от размеров узла, данный подход покажет наибольшую производительность при использовании в алгоритмах, принимающих на вход крупные блоки точек (например, для визуализации).



СПИСОК ЛИТЕРАТУРЫ

1. **Bi Z.M., Wang L.** Advances in 3D data acquisition and processing for industrial applications // *Robotics and Computer-Integrated Manufacturing*. 2010. Vol. 26. Pp. 403–413.
2. **Bellekens B., Spruyt V., Berkvens R., Weyn M.** A survey of rigid 3D point cloud registration algorithms // *AMBIENT 2014: the 4th Internat. Conf. on Ambient Computing, Applications, Services and Technologies*. 2014. Pp. 8–13.
3. **Nguyen A., Le B.** 3D point cloud segmentation: A survey // *The 6th IEEE Conf. on Robotics, Automation and Mechatronics*. 2013. Pp. 225–230.
4. **Grilli E., Menna F., Remondino F.** A review of point clouds segmentation and classification algorithms // *The Internat. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2017. Vol. 42. P. 339.
5. **Savage J.E.** Models of computation. Exploring the power of computing Addison-Wesley. Reading, MA. 1998.
6. **Liu X., Meng W., Guo J., Zhang X.** A survey on processing of large-scale 3D point cloud // *Internat. Conf. on Technologies for E-Learning and Digital Entertainment*. 2016. Pp. 267–279.
7. **Elseberg J., Borrmann D., N chter A.** One billion points in the cloud – An octree for efficient processing of 3D laser scans // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2013. Vol. 76. Pp. 76–88.
8. **Wang J., Lindenbergh R. Menenti, M.** SigVox – A 3D feature matching algorithm for automatic street object recognition in mobile laser scanning point clouds // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2017. Vol. 128. Pp. 111–129.
9. **Sveier A., Kleppe A.L., Tingelstad L., Egeland O.** Object detection in point clouds using conformal geometric algebra // *Advances in Applied Clifford Algebras*. 2017. Vol. 27. № 3. Pp. 1961–1976.
10. **Wang L., Xu Y., Li Y.** A Voxel-based 3D building detection algorithm for airborne LiDAR point clouds // *J. of the Indian Society of Remote Sensing*. 2019. Vol. 47. № 2. Pp. 349–358.
11. **Zhong L., Cheng L., Xu H., Wu Y., Chen Y., Li M.** Segmentation of individual trees from TLS and MLS Data // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2017. Vol. 10. № 2. Pp. 774–787.
12. **Xu Y., Tuttas S., Hoegner L., Stilla U.** Voxel-based segmentation of 3D point clouds from construction sites using a probabilistic connectivity model // *Pattern Recognition Letters*. 2018. Vol. 102. Pp. 67–74.
13. **Xu Y., Tuttas S., Hoegner L., Stilla U.** Geometric primitive extraction from point clouds of construction sites using VGS // *IEEE Geoscience and Remote Sensing Letters*. 2017. Vol. 14. № 3. Pp. 424–428.
14. **Tang Y., Feng J.** Multi-scale surface reconstruction based on a curvature-adaptive signed distance field // *Computers and Graphics (Per-gamon)*. 2018. Vol. 70. Pp. 28–38.
15. **Zhao R., Pang M., Liu C., Zhang Y.** Robust normal estimation for 3D LiDAR point clouds in urban environments // *Sensors (Switzerland)*. Vol. 19. No. 5.
16. **Farooqui M.Z., Shoaib M., Khan M.Z.** A comprehensive survey of page replacement algorithms // *IJARCET*. 2014.
17. **Fu S., He L., Huang C., Liao X.K., Li K.** Performance optimization for managing massive numbers of small files in distributed file systems // *IEEE Transactions on Parallel and Distributed Systems*. 2015. Vol. 26. Pp. 3433–3448.
18. **Samet H.** An overview of quadrees, octrees, and related hierarchical data structures // *Theoretical Foundations of Computer Graphics and CAD*. 1988. Pp. 51–68.
19. **Vo A.V., Laefer D.F., Bertolotto M.** Airborne laser scanning data storage and indexing: state-of-the-art review // *Internat. J. of Remote Sensing*. 2016. Vol. 37. Pp. 6187–6204.
20. **Photogrammetry & remote sensing A.S.P.R.S.** LAS specification version 1.4--R13 // URL: https://www.asprs.org/a/society/committees/standards/LAS_1_4_r13.pdf (Дата обращения: 08.10.2019).
21. **Rusu R.B., Cousins S.** Point cloud library (pcl) // *2011 IEEE Internat. Conf. on Robotics and Automation*. 2011. Pp. 1–4.
22. **Girardeau-Montaut D.** Cloud compare 3D point cloud and mesh processing software // *Open Source Project*. 2015.

Статья поступила в редакцию 10.11.2019.

REFERENCES

1. **Bi Z.M., Wang L.** Advances in 3D data acquisition and processing for industrial applications. *Robotics and Computer-Integrated Manufacturing*, 2010, Vol. 26, Pp. 403–413.
2. **Bellekens B., Spruyt V., Berkvens R., Weyn M.** A survey of rigid 3D point cloud registration algorithms. *AMBIENT 2014: the Fourth International Conference on Ambient Computing*,

Applications, Services and Technologies, 2014, Pp. 8–13.

3. **Nguyen A., Le B.** 3D point cloud segmentation: A survey. *The 6th IEEE Conference on Robotics, Automation and Mechatronics*, 2013, Pp. 225–230.

4. **Grilli E., Menna F., Remondino F.** A review of point clouds segmentation and classification algorithms. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, Vol. 42, P. 339.

5. **Savage J.E.** *Models of computation. Exploring the power of computing Addison-Wesley*. Reading, MA, 1998.

6. **Liu X., Meng W., Guo J., Zhang X.** A survey on processing of large-scale 3D point cloud. *International Conference on Technologies for E-Learning and Digital Entertainment*, 2016, Pp. 267–279.

7. **Elseberg J., Borrmann D., N chter A.** One billion points in the cloud – An octree for efficient processing of 3D laser scans. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, Vol. 76, Pp. 76–88.

8. **Wang J., Lindenbergh R. Menenti, M.** SigVox – A 3D feature matching algorithm for automatic street object recognition in mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, Vol. 128, Pp. 111–129.

9. **Sveier A., Kleppe A.L., Tingelstad L., Egeland O.** Object detection in point clouds using conformal geometric algebra. *Advances in Applied Clifford Algebras*, 2017, Vol. 27, No. 3, Pp. 1961–1976.

10. **Wang L., Xu Y., Li Y.** A Voxel-based 3D building detection algorithm for airborne LiDAR point clouds. *Journal of the Indian Society of Remote Sensing*, 2019, Vol. 47, No. 2, Pp. 349–358.

11. **Zhong L., Cheng L., Xu H., Wu Y., Chen Y., Li M.** Segmentation of individual trees from TLS and MLS Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, Vol. 10, No. 2, Pp. 774–787.

12. **Xu Y., Tuttas S., Hoegner L., Stilla U.** Voxel-based segmentation of 3D point clouds from

construction sites using a probabilistic connectivity model. *Pattern Recognition Letters*, 2018, Vol. 102, Pp. 67–74.

13. **Xu Y., Tuttas S., Hoegner L., Stilla U.** Geometric primitive extraction from point clouds of construction sites using VGS. *IEEE Geoscience and Remote Sensing Letters*, 2017, Vol. 14, No. 3, Pp. 424–428.

14. **Tang Y., Feng J.** Multi-scale surface reconstruction based on a curvature-adaptive signed distance field. *Computers and Graphics (Pergamon)*, 2018, Vol. 70, Pp. 28–38.

15. **Zhao R., Pang M., Liu C., Zhang Y.** Robust normal estimation for 3D LiDAR point clouds in urban environments. *Sensors (Switzerland)*, Vol. 19, No. 5.

16. **Farooqui M.Z., Shoaib M., Khan M.Z.** A comprehensive survey of page replacement algorithms, *IJARCET*, 2014.

17. **Fu S., He L., Huang C., Liao X.K., Li K.** Performance optimization for managing massive numbers of small files in distributed file systems. *IEEE Transactions on Parallel and Distributed Systems*, 2015, Vol. 26, Pp. 3433–3448.

18. **Samet H.** An overview of quadrees, octrees, and related hierarchical data structures. *Theoretical Foundations of Computer Graphics and CAD*, 1988, Pp. 51–68.

19. **Vo A.V., Laefer D.F., Bertolotto M.** Airborne laser scanning data storage and indexing: state-of-the-art review. *International Journal of Remote Sensing*, 2016, Vol. 37, Pp. 6187–6204.

20. Photogrammetry & Remote Sensing A.S.P.R.S. LAS Specification Version 1.4--R13. Available: https://www.asprs.org/a/society/committees/standards/LAS_1_4_r13.pdf (Accessed: 08.10.2019).

21. **Rusu R.B., Cousins S.** Point cloud library (pcl). *2011 IEEE International Conference on Robotics and Automation*, 2011, Pp. 1–4.

22. **Girardeau-Montaut D.** *Cloud compare 3D point cloud and mesh processing software. Open Source Project*, 2015.

Received 10.11.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

БЕЛЯЕВСКИЙ Кирилл Олегович

BELIAEVSKII Kirill O.

E-mail: kirill.beliaeviskii@spbpu.com

DOI: 10.18721/JCSTCS.12409
УДК 004.94

МЕТОД И ТЕХНОЛОГИЯ ИНТЕГРАЦИИ ДИСКРЕТНОЙ И СПЛОШНОСРЕДНОЙ МОДЕЛЕЙ ТРАНСПОРТНЫХ ПОТОКОВ РЕГИОНА

В.В. Глазунов¹, М.В. Чуватов¹, А.С. Чернышев², Л.М. Курочкин¹

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация;

² Физико-технический институт имени А.Ф. Иоффе РАН,
Санкт-Петербург, Российская Федерация

Рассмотрены методы построения континуальных макромоделей и дискретных микромоделей, описаны их ограничения. На основе дискретной модели реализованы сценарии движения транспортных средств в регионах. Рассмотрена континуальная макро модель MFlow собственной разработки, основой для которой являются уравнения Навье-Стокса, что позволяет использовать существующие наработки по численному решению уравнений сохранения и исследованию пузырьковых течений. Разработан метод интеграции дискретной и континуальной модели. Приведены параметры конфигурации и технология моделирования в дискретном симуляторе SUMO. Описаны параметры солвера MFlow и алгоритм решения уравнений в континуальной модели. Проведены эксперименты по сходимости результатов, получаемых в макро- и микромоделях для сценариев движения по прямым участкам дорог и Т-образным перекресткам при различной интенсивности движения транспортных средств.

Ключевые слова: континуальная модель, макро модель, микро модель, SUMO, TraCI, дискретный симулятор, сплошнoсредная модель, моделирование, транспортные потоки, интеллектуальные транспортные системы.

Ссылка при цитировании: Глазунов В.В., Чуватов М.В., Чернышев А.С., Курочкин Л.М. Метод и технология интеграции дискретной и сплошнoсредной моделей транспортных потоков региона // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 111–122. DOI: 10.18721/JCSTCS.12409

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

THE METHOD AND TECHNOLOGY FOR INTEGRATING DISCRETE AND CONTINUOUS MODELS OF REGIONAL TRAFFIC FLOWS

V.V. Glazunov¹, M.V. Chuvatov¹, A.S. Chernyshev², L.M. Kurochkin¹

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation;

² Ioffe Physical Technical Institute of the Russian Academy of Sciences,
St. Petersburg, Russian Federation

The paper considers the methods of constructing continual macromodels and discrete micromodels, their limitations are described. Based on the discrete model, scenarios of vehicle movement in the regions are implemented. The continual MFlow macromodel of its own design is considered, the basis for which are the Navier-Stokes equations, which allows

the use of existing developments in the numerical solution of conservation equations and the study of bubble flows. The method for integrating a discrete and continuum model is developed. The configuration parameters and modeling technology in the discrete SUMO simulator are presented. The parameters of the MFlow solver and the algorithm for solving equations in the continual model are described. Experiments were conducted on the convergence of the results obtained in macro and micromodels for scenarios of movement along straight sections of roads and T-shaped intersections at different traffic intensities.

Keywords: continual model, macromodel, micromodel, SUMO, TraCI, discrete simulator, continuous-medium model, modeling, traffic flows, intelligent transport systems.

Citation: Glazunov V.V., Chuvatov M.V., Chernyshev A.S., Kurochkin L.M. The method and technology for integrating discrete and continuous models of regional traffic flows. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 111–122. DOI: 10.18721/JCSTCS.12409

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Задачи моделирования транспортных потоков [1] решаются, начиная с 50-х гг. XIX века. Изначально для моделирования использовались макромодел, в дальнейшем развитие вычислительных средств позволило опуститься на уровень микро-моделирования, позволяющий учитывать особенности движения отдельных транспортных средств [2] и профили водителей [3]. Макромодел [4] дают возможность рассматривать потоки движения транспорта в целом и оперировать усредненными показателями, но не позволяют учитывать траектории движения каждой машины. Примерами макромоделей являются гидродинамические модели движения [5], в которых характеристики движения потока машин рассматриваются как процессы протекания жидкости или газа в трубе. Моделирование в микромоделах происходит посредством решения уравнений для движущихся объектов, согласно выбранной модели движения, например, следования за лидером [6].

Моделирование транспортных потоков дает возможность определять пропускную способность трасс и дорог в регионах и населенных пунктах, предсказывать образование заторов при проектировании дорожной сети. Процесс моделирования дорожного трафика облегчает оценку изменений

инфраструктуры до их реализации, например, эффективность зон охраны окружающей среды или алгоритмов управления светофорами, которые могут быть проверены и оптимизированы в симуляторе перед развертыванием в реальном мире.

Цель статьи – описание разработки механизма сопряжения макро- и микромоделей [7] на примере среды моделирования SUMO [8] и солвера MFlow [9]. Данная интеграция поможет уменьшить вычислительные затраты для расчета модели для дорожных участков с большой протяженностью, а также для участков с высокой плотностью дорожного трафика.

Дискретная модель SUMO

SUMO (Simulation of Urban Mobility) – это дискретно-событийный пакет моделирования движения транспортных средств с открытым исходным кодом, развивающийся с 2001 года. SUMO позволяет моделировать транспортные системы, включая легковые и грузовые транспортные средства, общественный транспорт и пешеходов [10]. В комплект SUMO входит набор инструментов, выполняющих такие задачи, как поиск маршрута, визуализация, создание транспортной сети. SUMO дает возможность создавать пользовательские модели и предоставляет API для удаленного управления симуляцией и интеграции.

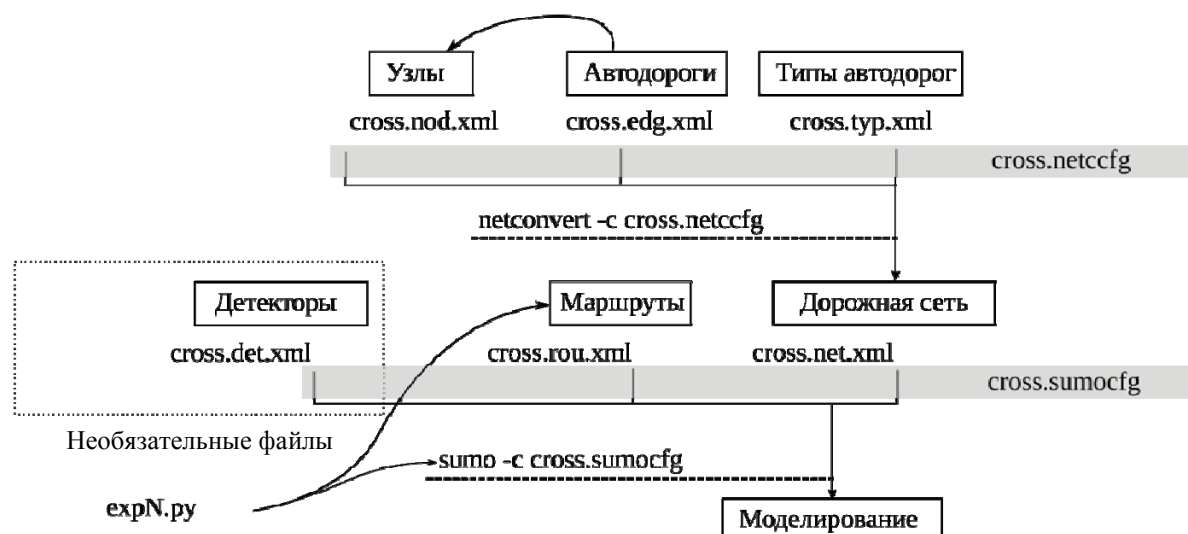


Рис. 1. Схема конфигурации модели при запуске моделирования через TraCI

Fig. 1. Model configuration when running simulation from TraCI

TraCI (Traffic Control Interface) – модуль симулятора, предоставляющий интерфейс API к моделированию дорожного движения, он позволяет извлекать значения моделируемых объектов и манипулировать их поведением в режиме реального времени. TraCI использует сетевую клиент-серверную архитектуру для взаимодействия с симулятором SUMO. Схема конфигурирования модели приведена на рис. 1.

Конфигурация описывает модель посредством определения параметров в xml-файлах. В файле `expN.py` указываются параметры запуска транспортных средств, им сопоставляются маршруты движения, задаются счетчики для контроля над процессом моделирования и указывается путь к конфигурационному файлу сценария `*.sumocfg`, в котором перечислены файлы описания дорожной сети `*.net.xml` и маршрутов движения транспортных средств `*.rou.xml`. Также возможно указать путь к файлу `*.trip.xml`, в который будут записываться треки движения транспортных средств в результате моделирования. Файл описания дорожной сети генерируется с помощью входящей в состав пакета SUMO утилиты `netconvert`, которая в качестве входных параметров принимает пути к файлам описания узлов дорожной сети `*.nod.xml`, соединяющих узлы автодорог `*.edg.xml`, типов автодорог `*.type.xml`, а выходным параметром указывается имя создаваемого файла описания дорожной сети.

Пример вызова утилиты `netconvert`:

```
netconvert.exe --node-files xxx.nod.xml --
edge-files yyy.edg.xml -t zzz.type.xml -o
www.net.xml
```

Файл описания узлов дорожной сети содержит перечень узлов в следующем формате: идентификатор узла, координата X (в метрах), координата Y (в метрах). Пример файла для прямой автодороги протяженностью 500 м:

```
<nodes>
  <node id="n01" x="-250" y="0" />
  <node id="n02" x="250" y="0" />
</nodes>
```

Файл описания автодорог содержит перечень участков дорог, соединяющих перечисленные в файле `*.nod.xml` узлы, в следующем формате: идентификатор узла начала участка, идентификатор узла конца участка, идентификатор участка, тип участка. Пример файла для прямой автодороги, соединяющей узлы `n01` и `n02`, из приведенного выше файла описания узлов:

```
<edges>
  <edge from="n01" to="n02" id="01to02"
type="2L60" />
</edges>
```

Файл описания типов автодорог содержит перечень используемых в эксперименте автодорог в следующем формате: идентификатор типа автодороги, приоритет дороги (используется для определения порядка движения транспортных средств на

нерегулируемых перекрестках), число рядов движения в одном направлении, максимально разрешенная скорость движения по данному типу дороги (м/с). Пример файла для прямой двухполосной дороги с максимальной скоростью 60,12 км/ч:

```
<types>
  <type id="2L60" priority="4" numLanes="2"
speed="16.7" />
</types>
```

Файл описания маршрутов движения транспортных средств содержит перечень типов транспортных средств и перечень маршрутов, по которым осуществляется движение в эксперименте. Типы транспортных средств описываются в следующем формате: ускорение при разгоне (м/с), ускорение при торможении (м/с), идентификатор типа транспортного средства, длина транспортного средства (в метрах), максимальная скорость транспортного средства (м/с), параметр отклонения характера движения от условно-идеального водителя (от 0.0 до 1.0). Маршруты движения транспортных средств описываются в следующем формате: идентификатор маршрута, перечень идентификаторов участков автодорог из файла *.edg.xml, которые должно проследовать транспортное средство в том порядке, в котором эти участки перечислены. Пример файла для эксперимента с движением легковых автомобилей по одному прямому участку дороги:

```
<routes>
  <vType accel="3.0" decel="5.0" id="Car"
length="4.5" maxSpeed="50" sigma="0" />
  <route id="r01" edges="01to02" />
</routes>
```

Архитектура среды моделирования, приведенная на рис. 2, обеспечивает интеграцию микро- и макромоделей с использованием сервера моделирования SUMO и подключение внешних сред моделирования посредством сетевых соединений.

Модуль junction позволяет установить на время симуляции счетчики, фиксирующие прохождение транспортных средств через них. Модуль vehicle позволяет генерировать объект транспортного средства на заданной дороге и полосе движения. Модуль output записывает результат срабатывания детекторов в CSV-файл или отправляет через внешний сокет во внешнюю модель, фиксирует реальное время компьютерного моделирования.

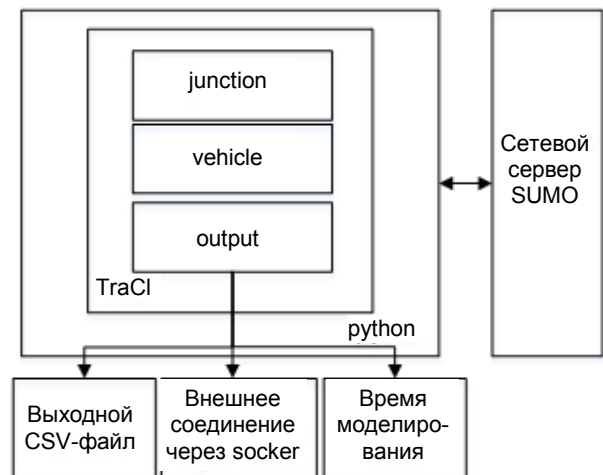


Рис. 2. Архитектура дискретной среды моделирования SUMO

Fig. 2. Architecture of SUMO discrete modelling

Континуальная макро модель MFlow

В основе континуальной модели дорожного трафика лежат уравнения сохранения плотности транспортного потока [11] и его скорости в каждой точке транспортной сети. Подход к описанию движения транспортных средств аналогичен сплошноредному подходу к описанию движения жидкости и газа по трубам, основополагающими уравнениями в котором являются уравнения Навье-Стокса [12]. Таким образом, возможно использование существующих наработок по численному решению уравнений сохранения [13], предложенных в работах по исследованию, в частности, пузырьковых течений [14].

Система уравнений континуальной модели дорожного трафика аналогична системе уравнений Навье-Стокса для сжимаемой среды:

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \frac{\partial(\rho V)}{\partial r} &= 0, \\ \frac{\partial \rho V}{\partial t} + V \frac{\partial(\rho V)}{\partial r} &= -\frac{\partial P}{\partial r} + \frac{\rho}{\tau}(V_e - V), \\ P &= \rho \cdot \theta.\end{aligned}$$

Здесь ρ — плотность транспортного потока, 1/м; V — скорость потока, м/с; P — эквивалентное давление, отвечающее за влияние соседних автомобилей, м/с²; τ — характерное время релаксации, с; V_e — равновесная

скорость потока, м/с; θ — модельный параметр, отвечающий за перемежаемость автомобилей в потоке и связь между давлением и плотностью, $\text{м}^2/\text{с}^2$.

Первое слагаемое в правой части уравнения сохранения импульса отвечает за взаимодействие соседних автомобилей и нарастает по мере увеличения плотности потока, что отражает взаимодействие отдельных водителей в микромоделях.

Второе слагаемое в правой части выражает стремление потока к достижению некоторой скорости V_e , являющейся оптимальной в данных условиях. Скорость V_e есть равновесная скорость потока, реализуемая при равномерном и гомогенном режиме движения по дороге и отражающая накладываемые ограничения на скоростной режим. Параметр τ является характерным временем релаксации потока на каждом участке дороги и в общем случае зависит от параметров потока и типа транспортного средства.

На входе в расчетную область, что соответствует началу участка дорожной сети, задаются фиксированные значения средней скорости и средней плотности потока. Значения параметров потока берутся из расчета по модели SUMO при помощи процедуры усреднения по времени. На выходе задаются мягкие граничные условия на скорость и плотность потока, а также фиксированное давление.

Для решения системы дифференциальных уравнений используется метод конечных объемов. При этом предполагается, что все величины заданы в центрах ячеек. Построение расчетной сетки осуществляется выдавливанием произвольного сечения вдоль направляющей линии дороги. Таким образом, на каждом участке сетка является квазиодномерной. Проинтегрированные уравнения сохранения для каждой из ячеек сетки имеют вид [15]:

$$\int_t^{t+\Delta t} \left(\frac{\partial}{\partial t} \int_{\Omega} \rho \phi d\Omega \right) dt + \int_t^{t+\Delta t} \int_{S_f} \rho \phi \vec{V} \cdot \vec{n} dS_f dt = \int_t^{t+\Delta t} \int_{\Omega} Q d\Omega dt,$$

где Ω — объем ячейки; S_f — площадь поверхности ячейки; f — грань ячейки; ϕ — искомая переменная; Δt — шаг по времени; Q — источниковое слагаемое.

При дискретизации по времени используется неявная схема первого порядка точности. Выражения для всех слагаемых записываются через неизвестные величины на новом временном шаге. Выбор такой схемы обусловлен тем, что она проста в реализации и позволяет выполнять вычисления с большими шагами по времени.

Полное уравнение сохранения с нестационарным слагаемым может использоваться и для расчета стационарных задач. Для обеспечения стабильности расчетной схемы шаг по времени выбирается исходя из критерия

$$\text{Куранта-Фридрихса-Леви } CFL = \Delta t \frac{U}{L}.$$

При выборе шага по времени параметр CFL не должен превосходить заданное значение. Здесь Δt — шаг по времени; U — масштаб скорости; L — линейный масштаб задачи. Для неявной схемы значения CFL могут превосходить единицу.

Выражения для вычисления поверхностного интеграла включают в себя неизвестные значения переменной ϕ_f на гранях ячеек. Для определения значений ϕ_f используются противопоточные схемы второго порядка точности.

Для обеспечения монотонности схема должна обладать свойством TVD (невозрастание полной вариации функции) [16]. В общем случае противопоточные схемы высоких порядков, обладающие свойством TVD, строятся по следующему принципу: в выражение для определения ϕ_f с помощью схемы первого порядка точности, обладающее большой численной диффузией, добавляется антидиффузионное слагаемое. Антидиффузионное слагаемое умножается на нелинейную функцию-ограничитель $\psi(r)$, зависящую от отношения градиентов величины вверх по потоку. Для схем второго порядка точности выражение для ϕ_f имеет вид:

$$\phi_f = \phi_P + \frac{1}{2} \psi(r) (\phi_D - \phi_P),$$

где $r = \frac{\phi_P - \phi_U}{\phi_D - \phi_P}$; P — центр ячейки, включающий грань f ; D — центр ячейки, расположенной вниз по потоку; U — центр ячейки, расположенной вверх по потоку.

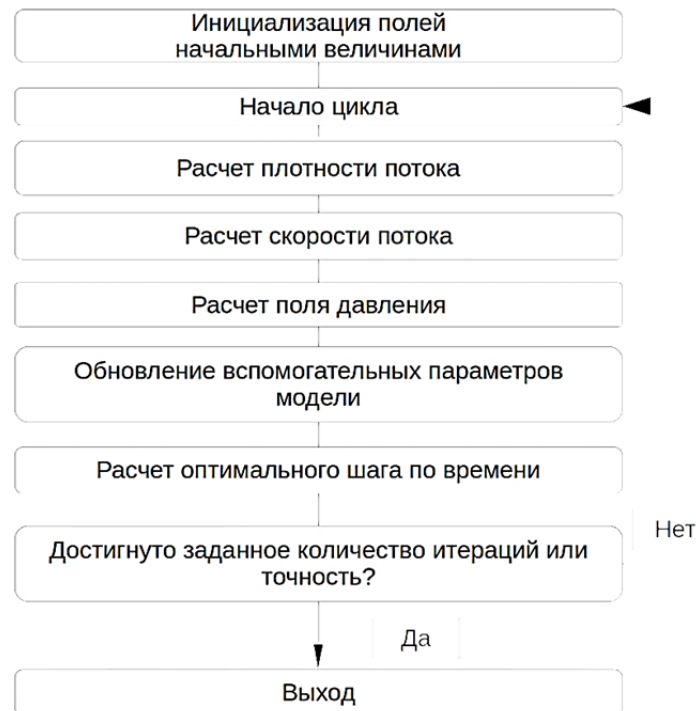


Рис. 3. Блок-схема решения уравнений
Fig. 3. Diagram for solving equations

В литературе существует большое количество выражений для ограничителя $\psi(r)$. В работе использованы три типа ограничителей: ограничитель ван Лира [17] — $\psi(r) = \frac{r+|r|}{1+|r|}$, ограничитель *minmod* — $\psi(r) = \max(0, \min(1, r))$ и ограничитель *superbee* — $\psi(r) = \max(0, \min(1, 2 \cdot r), \min(2, r))$ [18]. На основе данных, найденных в литературе [19], для проведения большинства расчетов был выбран ограничитель *minmod* как компромисс по совокупности точности и стабильности расчета.

В целях повышения устойчивости численной схемы используется линеаризация источниковых слагаемых. Для этого источниковое слагаемое разлагается в ряд Тейлора до первой производной относительно переменной ϕ :

$$Q = Q^k + \left(\frac{\partial Q}{\partial \phi} \right)^k (\phi - \phi^k).$$

Величины с индексом k известны и берутся с предыдущего шага по времени.

Это выражение можно переписать в следующем виде:

$$Q = Q_c + Q_p \phi, Q_c = Q^k - \left(\frac{\partial Q}{\partial \phi} \right)^k \phi^k, Q_p = \left(\frac{\partial Q}{\partial \phi} \right)^k.$$

Величина Q_c известна и остается в правой части уравнения. В том случае, когда $Q_p > 0$, перенос второго слагаемого в левую часть уравнения сохранения для ϕ приведет к увеличению диагонального преобладания результирующей матрицы и повышению стабильности численного метода.

Из-за существенной нелинейности решаемых уравнений необходимо ограничивать изменение переменных по ходу итерационного процесса. Обычно это достигается путем введения нижней релаксации. В этом случае значение величины на новой итерации зависит от величины на предыдущей итерации ϕ^* и приращения $\Delta\phi$ следующим образом: $\phi = \phi^* + \lambda \cdot \Delta\phi$.

Величина λ является параметром релаксации и выбирается из условия $0 < \lambda \leq 1$. Решение дискретизованных уравнений и обновление параметров течения происходит по схеме, приведенной на рис. 3, параметры солвера приведены в табл. 1.

Таблица 1
Параметры солвера MFlow

Table 1

MFlow solver parameters

Параметр	Значение
Точность линейных солверов	1E-06
Коэффициенты релаксации переменных	0.9–1.0
Количество итераций	10000
Куранта-Фридрихса-Леви (CFL)	1.0
Точность невязки для решаемых уравнений	1E-04

Интеграция моделей

Для интеграции моделей используется подход, основанный на том, что прямые участки и участки с большой плотностью автомобилей могут быть представлены в качестве сплошноредного потока транспортных средств в модели MFlow. Таким образом, используя архитектуру, представленную на рис. 2, можно подключить выход дорожных участков из дискретной модели в континуальную, а вход континуальной модели обратно в дискретную. Такой подход позволит выполнять моделирование выбранных участков параллельно. Формат выхода дискретной модели SUMO: time; start/end; vID; speed, где time – модельное время; start/end – признак завершения или начала движения ТС по дороге; vID – идентификатор ТС; speed – скорость ТС.

Формат выхода модели MFlow: массив из структур, содержащих time2; vID2; speed2, на основе выходных данных из MFlow запускаются новые машины в транспортном потоке на дискретной модели.

Постановка эксперимента

Для проверки результатов моделирования дискретной и непрерывной моделями, а также для оценки выигрыша в производительности подготовлены четыре сценария, реализованные в SUMO и в MFlow солвере. В первом сценарии (рис. 4 а) моделируется движение транспортных средств по прямому участку двухполосной односторонней автодороги протяженностью 500 м. Транспорт-

ные средства движутся на рисунке слева направо, иногда осуществляя перестроения с одной полосы на другую. Во втором сценарии моделируется движение транспортных средств по такой же дороге протяженностью 5000 м (рис. 4 а), также в направлении слева направо, осуществляя перестроения между полосами. В третьем сценарии (рис. 4 б) моделируется движение транспортных средств по трем однополосным односторонним дорогам протяженностью по 100 м, образующим Т-образный перекресток. Транспортные средства движутся слева направо до перекрестка, на котором в соотношении 1:1 осуществляют поворот налево или направо и продолжают движение, соответственно, вверх или вниз. В четвертом сценарии (рис. 4 б) моделируется движение транспортных средств по трем однополосным односторонним дорогам протяженностью по 100 м, также образующим Т-образный перекресток, но транспортные средства движутся сверху вниз до перекрестка, на котором в соотношении 1:3 осуществляют поворот направо или продолжают движение прямо.

Во всех сценариях для получения данных о количестве и скорости транспортных средств, а также о времени пересечения границ моделируемого участка, используются счетчики, установленные в конце автодорог. В первых двух сценариях используется два счетчика, в третьем и четвертом – три.

Начальная скорость транспортного средства при его появлении в модели равна максимально разрешенной, т. е. 16,7 м/с. Движение транспортного средства в целом отражает реальное поведение водителей. Общие значимые параметры моделей приведены в табл. 2.

В экспериментах с прямыми участками дороги протяженностью 500 и 5000 м транспортные средства запускаются поочередно на каждой из двух полос движения в единственном доступном направлении. При этом, как показало наблюдение за данным экспериментом, транспортные средства, движущиеся в левом ряду, стремятся набрать несколько более высокую скорость, чем движущиеся в правом ряду, что согласуется с поведением водителей в реальной дорожной обстановке. В экспериментах с Т-образным

Таблица 2

Параметры моделей SUMO и MFlow

Table 2

SUMO and MFlow model parameters

Параметр	Значение
Тип транспортного средства	Легковой автомобиль
Длина транспортного средства	4,5 м
Ширина транспортного средства	1,8 м
Максимальное ускорение и замедление трафика	3 м/с ²
Минимальный продольный интервал	2,5 м
Максимально разрешенная скорость движения по автодороге	16,7 м/с (60,12 км/ч)



Рис. 4. Схема движения: *a* – Прямой участок двухполосной односторонней автодороги протяженностью 500 или 5000 м; *б* – Т-образный перекресток с подъездом транспорта слева или сверху

Fig. 4. Traffic scheme: *a* – The straight section of a two-lane one-way road with a length of 500 or 5000 m; *б* – T-cross with vehicles flow from the left or top

перекрестком транспортные средства запускаются в заданном направлении по единственной доступной полосе движения. Если скорость движения транспортных средств и занятость участка дороги не позволяет разместить новое транспортное средство – его запуск отменяется до следующей попытки. Попытки запуска транспортных средств осуществляются каждые 2, 4, 10, 20, 40, 100 шагов симуляции. Запуск транспортных средств по числу шагов моделирования осуществляется для решения проблемы повторяемости эксперимента при многократных запусках. Как показало наблюдение за экспериментами, при одинаковых начальных условиях результаты моделирования в рамках выбранного сценария повторяются от запуска к запуску, что позволяет исключить случайные факторы при согласовании с непрерывной моделью.

В континуальной модели установлены соответствующие коэффициенты для запуска транспортных потоков, сопоставленные с дискретной моделью: 20 – 1; 40 – 0,5; 100 – 0,2; 10 – 2; 4 – 5; 2 – 10.

В результате проведенных экспериментов выявлены зависимости для сценариев движения транспортных средств, позволяющие использовать для расчета движущихся транспортных средств на определенных участках модель MFlow.

Так как модель SUMO дискретна, необходимо включить процедуру усреднения выходных данных модели для возможности их использования в качестве входных данных в континуальной модели дорожного трафика.

На рис. 5–7 темным цветом представлены графики изменения по времени мгновенной величины автомобильного потока на выходах из соответствующих дорожных

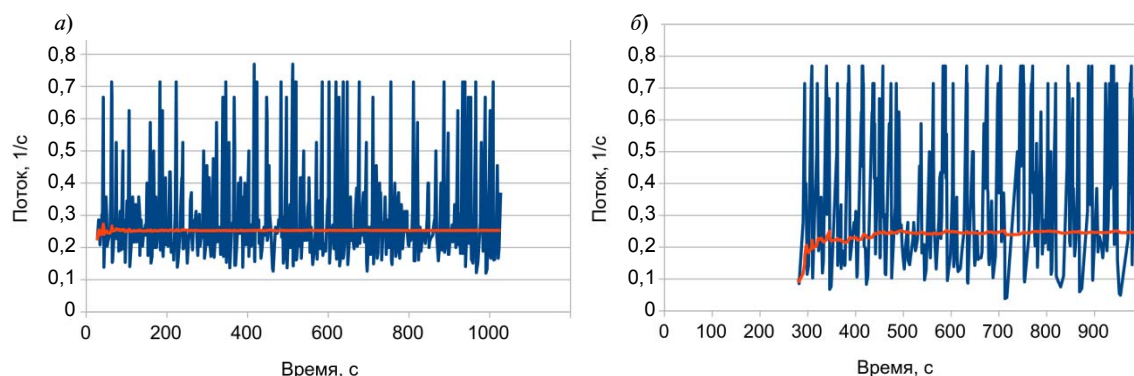


Рис. 5. Результаты моделирования SUMO и усредненные значения для прямого участка с интервалом запуска ТС 20 шагов: *a* – 500 м; *б* – 5000 м

(—) – усредненный транспортный поток; (—) – мгновенный транспортный поток

Fig. 5. SUMO simulation results and average values for the straight section with an interval of starting vehicles every 20 steps: *a* – 500 m; *б* – 5000 m

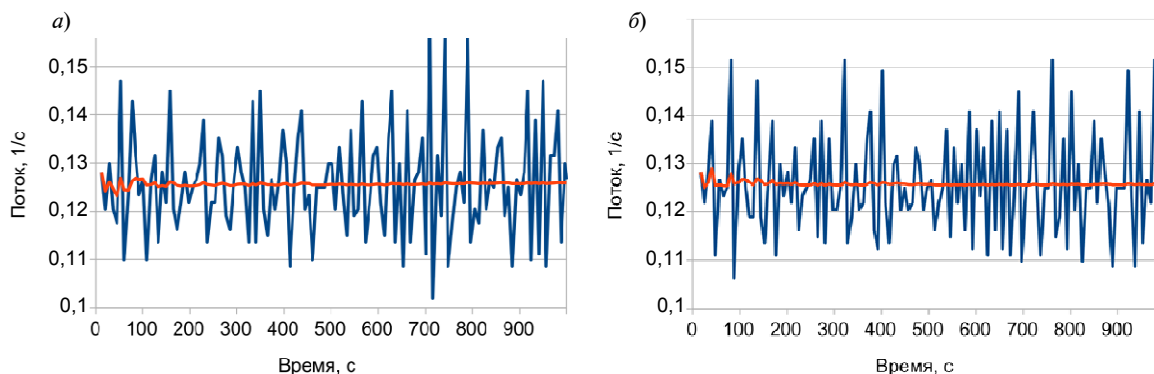


Рис. 6. Результаты моделирования SUMO и усредненные значения для Т-образного перекрестка с интервалом запуска ТС 40 шагов: *a* – поворот налево; *б* – поворот направо

(—) – усредненный транспортный поток; (—) – мгновенный транспортный поток

Fig. 6. SUMO simulation results and averaged values for the T-cross with an interval of starting vehicles every 40 steps: *a* – left turn; *б* – right turn

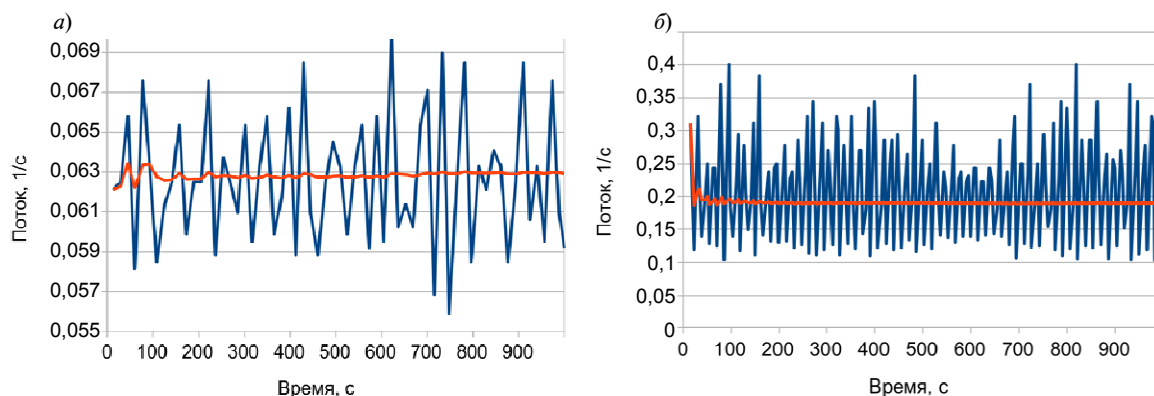


Рис. 7. Результаты моделирования SUMO и усредненные значения для Т-образного перекрестка с интервалом запуска ТС 40 шагов: *a* – поворот направо; *б* – движение прямо

(—) – усредненный транспортный поток; (—) – мгновенный транспортный поток

Fig. 7. SUMO simulation results and averaged values for the T-cross with an interval of starting vehicles every 40 steps: *a* – right turn; *б* – straight

участков, полученные по результатам работы модели SUMO. Несмотря на то, что входной поток машин задавался равномерным и односкоростным, результирующие значения существенно неоднородны по времени. Это связано с особенностью реализации модели, в которой каждый из участников дорожного движения подстраивается к особенностям потока за счет ускорения и замедления, а также перестроения или обгона других участников. Это приводит к отклонениям в скорости движения по мере прохождения участка дороги и, как следствие, к нестационарности поведения потока машин на выходе. Очевидно, что при заданном потоке машин на входе и отсутствии съездов (неконтролируемых источников и стоков для ТС), поток на выходе должен в стационарном режиме быть равен потоку на входе. Для определения выхода на стационарный режим использовалась процедура усреднения по времени, в качестве критерия достижения стационарности выступала нормированная на входной поток разность суммарных потоков на выходе и входе в область. Значение критерия не должно было превосходить некоторую заданную величину (в расчетах равную 0,001), достаточную для минимизации влияния неточности при расчете потока.

Усредненные по времени значения на рис. 5–7 представлены светлыми кривыми. Для каждого момента времени рассчитывалось среднее значение. С увеличением полного времени моделирования среднее значение стремилось к значению на входе. Отсчет среднего начинался с того момента, когда первая машина проходила выходную границу дороги. Как видно из анализа данных для прямых участков, модель SUMO быстро выходит на стационарный режим для случаев малой плотности потока или короткого участка. Это можно объяснить несколько меньшим разбросом по значениям для короткой дороги и малым влиянием соседних ТС для случая длинной дороги и малой плотности машин. Для

этих случаев выход на стационар занимал менее 800 с моделированного времени. С увеличением плотности время, необходимое для достижения нужной точности, растет и достигает 2000 с для варианта 5000 м с интервалом запуска ТС 20 шагов.

В случае с Т-образным перекрестком приведены данные по отдельным съездам. Видно, что когда потоки разделяются на два равных, сходимости у правого и левого съездов дороги примерно одинаковая. В том случае, когда потоки делятся в соотношении 1:3, сходимость у съезда с меньшим количеством машин худшая. Это объясняется тем, что на общем участке дороги накапливается неоднородность по потоку. После развилки из-за падения плотности необходимо большее время для накопления статистики, для того чтобы за счет достаточного количества машин выйти на требуемую точность.

Заключение

Произведено построение сетей дорожного трафика в макро- и микромоделях в типовых сценариях для прямых участков дорог и перекрестков с различными режимами движения. Результаты позволяют сделать вывод о применимости макроподхода для усредненных величин прохождения транспортного потока, рассчитываемых в макромоделях MFlow.

Разработано техническое решение, позволяющее построить гибридную модель движения транспортных средств, включающую сопряжение микро- и макромоделей, что позволяет использовать преимущество макромоделей в потоках высокой плотности и протяженности, получая константное время решения при таком же результате.

В дальнейшем возможно построение и сопоставление более сложных макромоделей, включающих наличие светофоров и дорожных препятствий на полосах.

Работа выполнена при финансовой поддержке РФФИ, проект №18-07-00430.

СПИСОК ЛИТЕРАТУРЫ

1. Швецов В.И., Алиев А.С. Математическое моделирование загрузки транспортных сетей. М.: Едиториал УРСС, 2003. 64 с.
2. Newell G.F. A simplified car-following theory: a lower order model. // *Transp. Res.* 2002. Vol. 36. Pp. 195–205.
3. Yang Q., Koutsopoulos H.N. A microscopic traffic simulator for evaluation of dynamic traffic management systems // *Transportation Research. Part C*. 1996. Vol. 4(3). Pp. 113–129.
4. Гасников А.В., Кленов С.Л., Нурминский Е.А., Холодов Я.А., Шамрай Н.Б. Введение в математическое моделирование транспортных потоков. М.: МФТИ, 2010. 362 с.
5. Krauß S., Wagner P., Gawron C. Metastable states in a microscopic model of traffic flow // *Physical Review E*. 1997. Vol. 55. No. 304. Pp. 55–97.
6. Hidas P. Modeling lane changing and merging in microscopic traffic simulation // *Transportation Research. Part C*. 2002. Vol. 10(5). Pp. 351–371.
7. Курочкин Л.М., Чуватов М.В., Глазунов В.В., Чернышев А.С. Сравнение результатов моделирования транспортных потоков сплошноредными и дискретно-событийными методами // Международная конф. по мягким вычислениям и измерениям. 2018. Т. 1. С. 500–503.
8. Haddouch S., Nachimi H., Hmina N. Modeling the flow of road traffic with the SUMO simulator // 4th Internat. Conf. on Optimization and Applications. Mohammedia. 2018. Pp. 1–5.
9. Чернышев А.С., Шмидт А.А. Использование эйлерово-эйлеровского подхода для моделирования турбулентных течений пузырьковых сред // Письма в журнал технической физики. 2013. Т. 39. № 12. С. 17–24.
10. Lim K.G., Lee C.H., Chin R.K.Y., Beng Yeo K., Teo K.T.K. SUMO enhancement for vehicular AD Hoc network (VANET) simulation // 2017 IEEE 2nd Internat. Conf. on Automatic Control and Intelligent Systems. Kota Kinabalu, 2017. Pp. 86–91.
11. Helbing D., Hennecke A., Shvetsov V., Treiber M. Micro- and macro- simulation of free-way traffic // *Mathematical and Computer Modelling*. 2002. Vol. 35. Pp. 517–547.
12. Helbing D. Gas-kinetic derivation of Navier-Stokes-like traffic equations // *Phys. Rev. E*. 1996. Vol. 53. Pp. 2366–2381.
13. Лойцянский Л.Г. Механика жидкости и газа. М.: Дрофа, 2006.
14. Chernyshev A., Schmidt A., Kurochkin L. Numerical modeling of polydisperse bubbly flows by the openMP parallel algorithm // *Proc. Comp. Sci. ICCS-2017. Zurich Switzerland*, 2017. Vol. 108C. Pp. 1990–1997.
15. Versteeg H.K., Malalasekera W. An introduction to computational fluid dynamics. Pearson Education Limited, UK, 2007.
16. Hirsch Ch. Numerical computation of internal & external flows: fundamentals of numerical discretization. New York: John Wiley & Sons, 1988.
17. van Leer B. Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme // *J. Comp. Phys.* 1974. Vol. 14. Pp. 361–370.
18. Roe P.L. Characteristic-based schemes for the Euler equations // *Ann. Rev. Fluid Mech.* 1986. Vol. 18. Pp. 337–365.
19. Чирков Д.В., Черный С.Г. Сравнение точности и сходимости некоторых TVD-схем // Вычислительные технологии. 2000. Т. 5. № 5. С. 86–107.

Статья поступила в редакцию 27.09.2019.

REFERENCES

1. Shvetsov V.I., Aliyev A.S. *Matematicheskoye modelirovaniye zagruzki transportnykh setey* [Mathematical modeling of loading transport networks]. Moscow: Yeditorial URSS Publ., 2003, 64 p. (rus)
2. Newell G.F. A simplified car-following theory: A lower order model. *Transp. Res.*, 2002, Vol. 36, Pp. 195–205.
3. Yang Q., Koutsopoulos H.N. A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research. Part C*, 1996, Vol. 4(3), Pp. 113–129
4. Gasnikov A.V., Klenov S.L., Nurminskiy Ye.A., Kholodov Ya.A., Shamray N.B. *Vvedeniye v matematicheskoye modelirovaniye transportnykh potokov* [Introduction to mathematical modeling of traffic flows]. Moscow: MFTI Publ., 2010, 362 p. (rus)
5. Krauß S., Wagner P., Gawron C. Metastable states in a microscopic model of traffic flow. *Physical Review E*, 1997, Vol. 55, No. 304, Pp. 55–97.
6. Hidas P. Modeling lane changing and merging in microscopic traffic simulation. *Transportation Research. Part C*, 2002, Vol. 10(5), Pp. 351–371.

7. **Kurochkin L.M., Chuvatov M.V., Glazunov V.V., Chernyshev A.S.** Sravneniye rezultatov modelirovaniya transportnykh potokov sploshnosrednymi i diskretno-sobytiynym metodami [Comparison of the results of modeling traffic flows with continuous and discrete-event methods]. *XXI Mezhdunarodnaya konferentsiya po myagkim vychisleniyam i izmereniyam [The XXI International Conference on Soft Computing and Measurement]*, 2018, Vol. 1, Pp. 500–503. (rus)
8. **Haddouch S., Hachimi H., Hmina N.** Modeling the flow of road traffic with the SUMO simulator. *2018 4th International Conference on Optimization and Applications*, Mohammedia, 2018, Pp. 1–5.
9. **Chernyshev A.S., Shmidt A.A.** Ispolzovaniye eylerovo-eylerovskogo podkhoda dlya modelirovaniya turbulentnykh techeniy puzyrkovykh sred [Using the Eulerian-Eulerian approach for modeling turbulent flows of bubble media]. *Pisma v Zhurnal Tekhnicheskoy Fiziki [Technical Physics Letters]*, 2013, Vol. 39, No. 12, Pp. 17–24. (rus)
10. **Lim K.G., Lee C.H., Chin R.K.Y., Beng Yeo K., Teo K.T.K.** SUMO enhancement for vehicular ad hoc network (VANET) simulation. *2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems*, Kota Kinabalu, 2017, Pp. 86–91.
11. **Helbing D., Hennecke A., Shvetsov V., Treiber M.** Micro- and macro- simulation of freeway traffic. *Mathematical and Computer Modelling*, 2002, Vol. 35, Pp. 517–547.
12. **Helbing D.** Gas-kinetic derivation of Navier-Stokes-like traffic equations, *Phys. Rev. E*, 1996, Vol. 53, Pp. 2366–2381.
13. **Loytsyanskiy L.G.** *Mekhanika zhidkosti i gaza [Fluid and gas mechanics]*. Moscow: Drofa Publ., 2006. (rus)
14. **Chernyshev A., Schmidt A., Kurochkin L.** Numerical modeling of polydisperse bubbly flows by the openMP parallel algorithm. *Proc. Comp. Sci. ICCS-2017, Zurich Switzerland, 2017*, Vol. 108C, Pp. 1990–1997.
15. **Versteeg H.K., Malalasekera W.** *An introduction to computational fluid dynamics*. Pearson Education Limited, UK, 2007.
16. **Hirsch Ch.** *Numerical computation of internal & external flows: fundamentals of numerical discretization*. New York: John Wiley & Sons, USA, 1988.
17. **van Leer B.** Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme. *J. Comp. Phys.*, 1974, Vol. 14, Pp. 361–370.
18. **Roe P.L.** Characteristic-based schemes for the Euler equations. *Ann. Rev. Fluid Mech.*, 1986, Vol. 18, Pp. 337–365.
19. **Chirkov D.V., Chernyy S.G.** Sravneniye tochnosti i skhodimosti nekotorykh TVD-skhem [Comparison of accuracy and convergence of some TVD schemes]. *Vychislitelnyye tekhnologii [Computational Technologies]*, 2000, Vol. 5, No. 5, Pp. 86–107. (rus)

Received 27.09.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

ГЛАЗУНОВ Вадим Валерьевич

GLAZUNOV Vadim V.

E-mail: neweagle@gmail.com

ЧУВАТОВ Михаил Владимирович

CHUVATOV Mikhail V.

E-mail: misha@iktp.spbstu.ru

ЧЕРНЫШЕВ Александр Сергеевич

CHERNYSHEV Alexander S.

E-mail: alexander.tchernyshev@mail.ioffe.ru

КУРОЧКИН Леонид Михайлович

KUROCHKIN Leonid M.

E-mail: kurochkinl@spbstu.ru

DOI: 10.18721/JCSTCS.12410
УДК 681.5

РАСЧЕТ ТОЧЕК ТЕХНОЛОГИЧЕСКОЙ ТРАЕКТОРИИ ЛАЗЕРНОГО РОБОТИЗИРОВАННОГО КОМПЛЕКСА ОТНОСИТЕЛЬНО ИНСТРУМЕНТА И ОСНАСТКИ В САМ

Д.В. Родионов, А.Б. Люхтер, В.Г. Прокошев

Владимирский государственный университет имени А.Г. и Н.Г. Столетовых,
Владимир, Российская Федерация

Рассмотрены методы расчета точек технологической траектории лазерного роботизированного комплекса (ЛРК) в мировом графическом пространстве САМ относительно инструмента и подвижной оснастки. Изучены вопросы использования САМ-систем в современных лазерных роботизированных комплексах, а также представлена программная разработка. Рассмотрены основные специфики и проблемы расчета точек мирового графического пространства САМ и их соответствие расположению инструмента ЛРК и подвижной оснастки. Представлены методы расчета положения и ориентации инструмента относительно локального базиса модели детали. Предложен их расчет в системе двухосевого позиционирующего устройства ЛРК. Описан механизм определения положения угла вращения вокруг лазерного луча исполнительного механизма ЛРК, закрепленного на шестой оси подвижного фланца последнего звена промышленного робота.

Ключевые слова: CAD/CAM, САПР, система проектирования, лазерные роботизированные комплексы, лазерная сварка, промышленные роботы, промышленные лазеры, цифровизация промышленности, оффлайн программирование.

Ссылка при цитировании: Родионов Д.В., Люхтер А.Б., Прокошев В.Г. Расчет точек технологической траектории лазерного роботизированного комплекса относительно инструмента и оснастки в САМ // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 123–135. DOI: 10.18721/JCSTCS.12410

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

CALCULATION OF THE POINTS OF THE TECHNOLOGICAL TRAJECTORY OF THE LASER ROBOTIC COMPLEX RELATIVE TO THE TOOL AND TOOLING IN THE CAM

D.V. Rodionov, A.B. Lyukhter, V.G. Prokoshev

Vladimir State University name after Alexander and Nikolay Stoletovs,
Vladimir, Russian Federation

The article deals with the methods of calculating the points of the technological trajectory of the laser robotic complex (LRC) in the world graphic space CAM relative to the tool and mobile equipment. The questions of using CAM systems in modern laser

robotic systems are considered, as well as the authors own software development is presented. The main specifics and problems of calculation of points of the world graphic space CAM and their correspondence to the location of the LRC tool and mobile equipment are considered. Methods for calculating the position and orientation of the tool relative to the local basis of the part model are presented. Their calculation in the system of two-axis positioning device of LRC is offered. The mechanism of determining the position of the rotation angle around the laser beam of the LRC actuator fixed on the 6th axis of the movable flange of the last link of the industrial robot is described.

Keywords: CAD/CAM, design system, laser robotic systems, laser welding, industrial robots, industrial lasers, industry digitalization, offline programming.

Citation: Rodionov D.V., Lyukhter A.B., Prokoshev V.G. Calculation of the points of the technological trajectory of the laser robotic complex relative to the tool and tooling in the CAM. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 123–135. DOI: 10.18721/JCSTCS.12410

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Использование компактных итербиевых волоконных лазеров на современных промышленных предприятиях подразумевает получение немалых преимуществ [1] и значительного расширения технических и конструкторских решений [2, 3].

В настоящее время современные комплексы [4] представляют собой объединения множества технических устройств, взаимодействующих на различных уровнях, с разной степенью отклика.

Использование систем автоматизированного проектирования в промышленном производстве является неотъемлемым тре-

бованием современного предприятия. Благодаря их применению технологии имеют возможность осуществлять цифровое проектирование процесса технологической обработки детали [5, 6]. На основе результата проектирования производится генерация управляющих программ аппаратных средств промышленного комплекса [7–9].

С проблемой использования САМ в обработке металлоизделий лазером на основе роботизированного манипулятора столкнулся коллектив авторов при автоматизации одной из своих разработок — лазерного роботизированного комплекса (ЛРК) (рис. 1).



Рис. 1. Схема лазерного роботизированного комплекса

Fig. 1. Scheme of laser robotic complex

Для выполнения требований автоматизации, предъявляемых современным промышленным предприятиям в эпоху четвертой промышленной революции [10–13], для ЛРК было решено использовать САМ-систему [14–16], которая в свою очередь должна обеспечивать следующее:

- быстрое цифровое проектирование процесса обработки как единичных, так и серийных изделий;
- моделирование движения подвижных элементов комплекса;
- обеспечение технологического и конструкторского расчета при проектировании;
- использование базы знаний для обработки конкретной детали;
- подготовку результата проектирования к предпроцессу (процессу контроля, выполняемому перед технологической обработкой);
- генерацию управляющих программ всех технических элементов промышленного комплекса, учитывая особенности возможной замены и расширения технических компонентов комплексов.

Нами был проведен анализ существующих систем автоматизированного проектирования, пригодных для использования ЛРК. Результат анализа показал, что на текущий момент существующие САМ-системы неполным образом удовлетворяют требованиям современной лазерной роботизированной обработки. Это связано с отсутствием:

- 1) возможности генерации управляющих программ технических устройств ЛРК (например, иттербиевого волоконного лазера);
- 2) возможности использования результата проектирования к подготовке программ пред- и постпроцессов (первичная корректировка, анализ результата и т. д.) лазерной роботизированной обработки изделий из металла;
- 3) возможности внедрения алгоритмов автоматизации (например, проектирование участков разгона/торможения инструмента и, как следствие, задание управляющих команд переключения сигналов оборудования по времени) лазерных технологических процессов при проектировании;

4) возможности использования базы знаний технологических режимов лазерной роботизированной обработки;

5) методов проектирования оптимальных технологических траекторий движения инструмента лазерного роботизированного комплекса;

6) методов для гибкой связи с системами контроля и управления ЛРК.

В связи с этим нами была произведена разработка и внедрение собственного прикладного САМ-продукта (рис. 2).

В ходе работы по созданию САМ-системы была выявлена необходимость в разработке методов проектирования и расчетов позиционирования *технологической траектории*¹ ЛРК в мировом графическом пространстве. Настоящая статья посвящена решению одной из задач, связанных с расчетом положения точек технологической траектории в локальном базисе модели.

Описание проблемы исследования

Исходными данными для работы с САМ является трехмерная полигональная модель детали [17]. Методы проектирования траектории на модели описаны нами в [18, 19].

На поверхности модели технолог осуществляет проектирование технологической траектории ЛРК, причем траектория в таком случае будет представлять собой последовательность точек с заданными технологическими значениями. В качестве примера параметров технологической траектории может служить иллюстрация параметров лазерной головки с навесным оборудованием (далее — инструмент) на рис. 3.

Большинство технологических параметров необходимы для генерации управляющих программ, и они подбираются технологом из базы данных режимов в зависимости от процесса и условий лазерной обработки, но такие параметры, как положение и ориентация инструмента, положение двухосевого позиционирующего устройства (далее — позиционер) требуют конструкторского подхода в системе САМ.

¹ Траектория движения ЛРК с заданными технологическими параметрами.

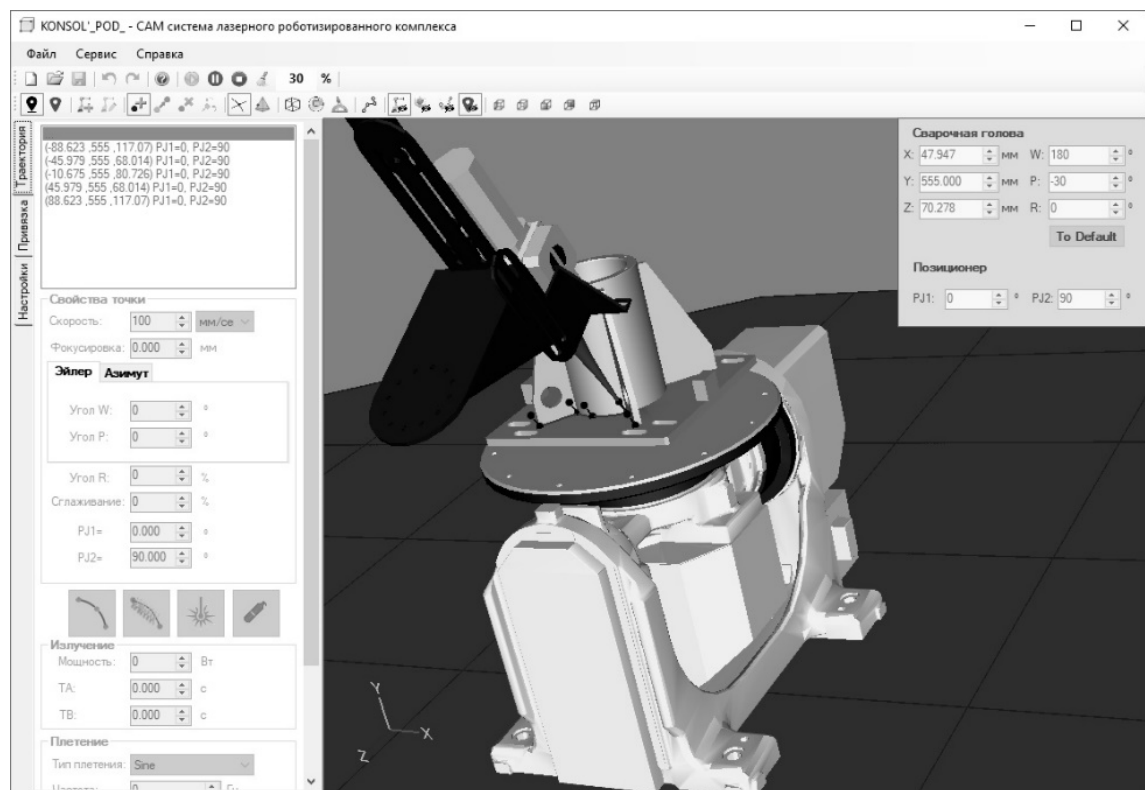


Рис. 2. CAM-система лазерного роботизированного комплекса

Fig. 2. CAM-system of laser robotic complex

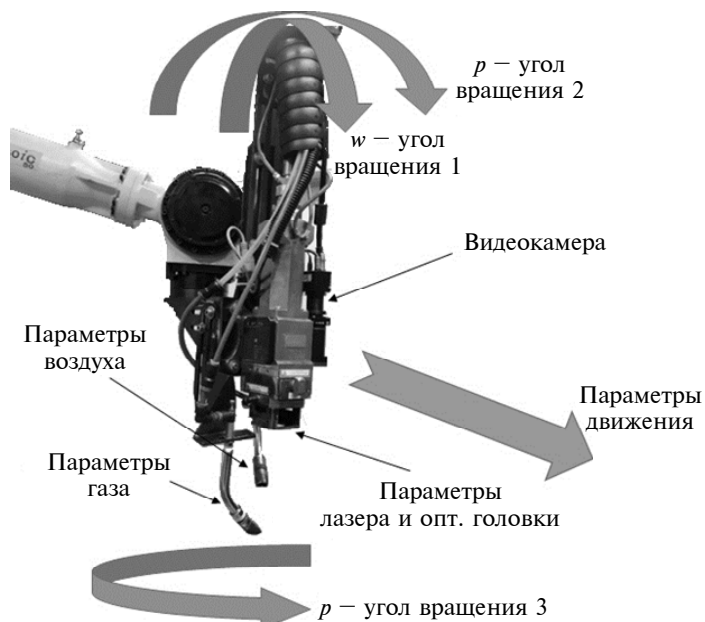


Рис. 3. Технологические параметры инструмента лазерного роботизированного комплекса

Fig. 3. The technological parameters of the laser tool of the robotic complex

Механизмы проектирования технологической траектории, разработанные и описанные нами в работе [18], не учитывают требования построения точек траектории вне поверхности детали (узлов графа полигональной модели).

Для решения задачи построения точек траектории в мировом графическом САМ-пространстве мы решили расширить функционал подсистемы моделирования движения инструмента и позиционера (подсистемы, позволяющей воспроизводить кинематику и динамику движения для спроектированной траектории), в частности, производить построение точки технологической траектории относительно инструмента ЛРК в локальном базисе модели [26].

При этом возникают следующие проблемы:

1. Точки технологической траектории представлены в локальном базисе модели, а инструмент осуществляет свое позиционирование в мировом графическом базисе.

2. Ориентация инструмента в каждой точке технологической траектории представляется в виде нормированного вектора, указывающего положение лазерного луча, а позиционирование инструмента в мировом графическом пространстве происходит в системе углов Эйлера (описывающих кинематику вращения инструмента ЛРК).

3. При переносе модели в рабочее пространство ЛРК (генерации управляющих программ), расположенной на позиционере, привязка² осуществляется в фиксированном положении позиционера. Деталь в ходе выполнения техпроцесса за счет подвижной оснастки (позиционера) может менять свое положение. Это в свою очередь требует пересчета положения точек траектории и ориентации инструмента при положениях детали, не соответствующих положению привязки.

² Привязка — процесс синхронизации детали и ее модели посредством применения математических методов расчета.

4. Для однозначного соответствия инструмента точке технологической траектории необходимо произвести расчет положения угла вращения вокруг лазерного луча инструмента ЛРК, который с технологической точки зрения должен иметь вполне конкретное значение.

Постановка задачи

Цель исследования — разработка метода расчета точек технологической траектории в мировом графическом САМ-пространстве лазерного роботизированного комплекса относительно инструмента в локальном базисе модели с учетом расположения детали в системе подвижной оснастки.

Пусть $\bar{p} = \{p_x, p_y, p_z\}$ — точка позиционирования³ (далее — TCP (Tool Center Point) (рис. 4)), $\bar{\gamma} = \{w, p, r\}$ — ориентация инструмента⁴, а J_1 и J_2 — углы вращения позиционера вокруг его осей j_1 и j_2 соответственно с нейтральным положением в осях j_1^0 и j_2^0 (рис. 5), тогда $\bar{p}'$ — точка TCP в локальном базисе модели, а $\bar{n} = \{n_x, n_y, n_z\}$ — нормированный вектор ориентации инструмента в точке технологической траектории. При такой постановке инструмент будет иметь нейтральное положение в значениях углов Эйлера $\bar{\gamma}_0 = \{180, 0, 0\}$, а позиционер будет находиться в нейтральном положении при $J_1 = 0, J_2 = 0$.

Мы производили расчет угла вращения r инструмента вокруг лазерного луча в зависимости от направления движения $\bar{d}$ вдоль обрабатываемой поверхности. Необходимость в таком расчете связана с требованием определенного и технологически верного расположения навесного оборудования (сопла газа и воздуха, триангуляционные датчики и т. д.) на инструменте в каждом отрезке зоны обработки.

³ В качестве точки позиционирования инструмента ЛРК выступает фокус лазерного излучения.

⁴ Ориентация представляет собой систему углов Эйлера.

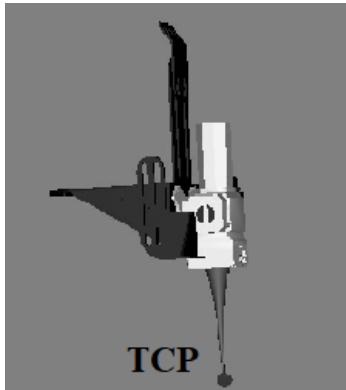


Рис. 4. Модель инструмента ЛРК с TCP в точке фокуса лазера

Fig. 4. Model tool of LRC with TCP at the laser focus point

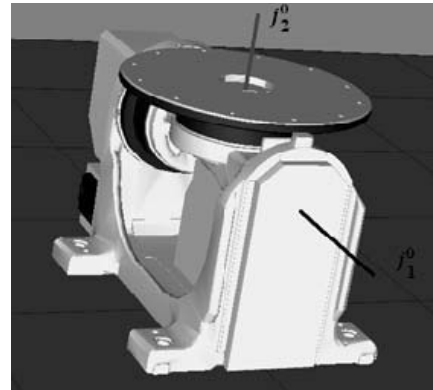


Рис. 5. Модель позиционера ЛРК с осями поворота в нулевом значении

Fig. 5. Model positioner of LRC with axes of rotation in the zero value

При построении точек относительно инструмента очевидным образом возникают ограничения, связанные с кинематикой движения инструмента и позиционера. Помимо этого авторы вносят дополнительные ограничения на область допустимых значений позиционирования главных подвижных элементов⁵ ЛРК, связанные с особенностями монтажа и использования оборудования. Все эти ограничения можно представить в следующем виде:

$$w \in [90, 180] \cup [-90, -180]; \quad (1)$$

$$p \in [-90, 90]; \quad (2)$$

$$r \in [-135, 135]; \quad (3)$$

$$J_1 \in [-135, 135]; \quad (4)$$

$$J_2 \in [-180, 180]. \quad (5)$$

Положения модели инструмента на основе которых были построены точки технологической траектории в САМ должны полностью соответствовать положениям исполнительного звена после переноса модели в рабочее пространство ЛРК. Выполнение этого условия можно использовать как критерий качества построенного решения.

В таком случае, для построения точки технологической траектории в мировом графическом пространстве относительно инструмента и подвижной оснастки ЛРК необходимо решить следующие задачи.

1. Представить положение и ориентацию инструмента в локальном базисе модели, т. е. найти преобразование $\vec{p} \rightarrow \vec{p}'$ и $\vec{\gamma} \rightarrow \vec{\gamma}'$.

⁵ Под главными подвижными элементами ЛРК понимается инструмент и позиционер.

2. Рассчитать значения $\vec{p}'$ и $\vec{\gamma}'$, в случае, когда модель детали располагается в системе подвижной оснастки (позиционера), используемой в ЛРК. При расчете учесть положение позиционера, при котором предполагается осуществлять привязку модели и детали.

3. Определить зависимости угла вращения инструмента r вокруг лазерного луча в зависимости от направления движения $\vec{d}$ инструмента вдоль зоны обработки.

Решения подобного рода задач представлены в различных источниках [20–23], но материал, описанный в них, не полностью применим для решения задач настоящего исследования, т. к. расчет позиционирования и ориентации исполнительного узла ЛРК должен быть представлен в локальном базисе модели. Специфика этого требования связана с особенностями программной реализации современной САМ-системы.

Представление инструмента ЛРК в локальном базисе модели

В качестве базиса мирового графического пространства САМ взят стандартный графический декартовый базис, определяемый следующими направлениями при взгляде на пространство, не подвергнутое видовому изменению⁶: ОХ — влево, ОУ — вверх, ОZ — на наблюдателя.

⁶ Изменения, подразумевающие преобразование исходного базиса (параллельный перенос, вращение и т. д.).

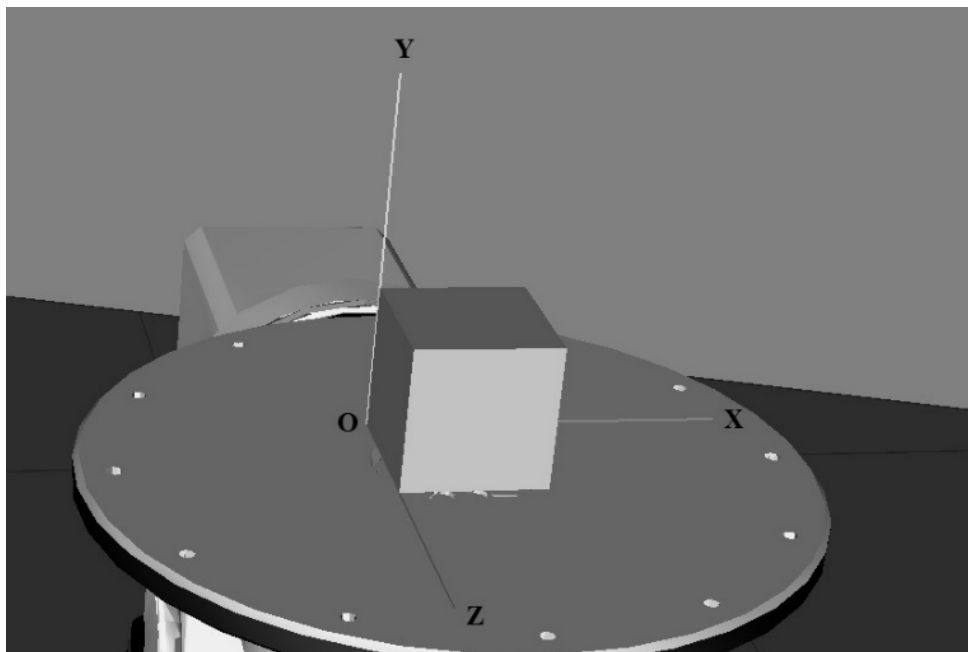


Рис. 6. Локальный базис модели куба, спроектированный относительно одной из его вершин (точка O)

Fig. 6. A local basis of a cube model designed with respect to one of its vertices (point O)

Направления базисных орт локального базиса модели (без видового изменения) совпадают с направлениями орт мирового графического пространства. В качестве начала отсчета локального базиса была выбрана точка нулевого значения координат модели. Полученный репер будем называть *локальным базисом модели* (рис. 6).

Для расчета положения инструмента $\bar{p}$ в локальном базисе модели стоит учитывать, что модель может быть смещена и повернута относительно исходного положения. Вращение модели детали в САМ происходит вокруг осей i и j (орт OX и OY соответственно). В таком случае положение инструмента $\bar{p}'$ относительно модели примет следующий вид:

$$\bar{p}' = \bar{p} - \bar{m}; \quad (6)$$

$$\bar{p}' = M(i, \alpha) \bar{p}'; \quad (7)$$

$$\bar{p}' = \bar{p}' + 2\bar{m}; \quad (8)$$

$$\bar{p}' = M(j, \beta) \bar{p}'; \quad (9)$$

$$\bar{p}' = \bar{p}' - 2\bar{m}, \quad (10)$$

где $M(\bar{v}, \varphi)$ — матрица поворота вокруг единичного вектора $\bar{v}$ на угол φ .

В выражениях (6), (8) и (10) вектор $\bar{m}$ характеризует текущее расположение модели в мировом графическом пространстве. Выражения (7) и (9) характеризуют соответствие инструмента и модели детали, осуществляющей поворот на углы α и β относительно осей i и j соответственно.

Углы ориентации инструмента ЛРК представляются в виде углов крена, тангажа и «рысканья» инструмента (повороты вокруг координатных осей OX, OY и OZ), обозначим их как w , p и r соответственно. Данная система определения ориентации в робототехнике известна, как система углов Эйлера [23, 24].

Рассмотрим вектор $\bar{n}_0 = \{0, -1, 0\}$, соответствующий значению углов Эйлера $\bar{\gamma}_0 = \{180, 0, 0\}$, нейтрального положения инструмента. Пусть мы хотим произвести поворот в системе углов Эйлера на значение $\bar{\gamma} = \{w, p, r\}$ и для него найти радиус-вектор $\bar{n}$. Для этого определим оси вращения $\{I, J, K\}$ в локальном базисе модели для углов $\bar{\gamma}$. В таком случае, несложно заметить, что для значений $\bar{n}_0$, $\bar{\gamma}_0$ оси вращения углов

$\bar{\gamma}$ выглядят следующим образом: $i = -OX$, $j = -OY$ и $k = -OZ$. Следовательно, для определения осей углов $\bar{\gamma}$ необходимо произвести следующие преобразования:

$$K = k; \quad (11)$$

$$I = M(K, w)i; \quad (12)$$

$$J = M(I, p)M(K, w)j. \quad (13)$$

Из (11)–(13) следует, что для получения радиус-вектора $\bar{n}$ необходимо осуществить последовательные повороты:

$$\bar{n} = M(J, r)M(I, p)M(K, w)\bar{n}_0. \quad (14)$$

Вектор $\bar{n}$ будет представлять собой нормированный вектор ориентации инструмента ЛРК в локальном базисе модели детали, соответствующий заданным углам Эйлера $\bar{\gamma}$.

Положение и ориентация инструмента в системе подвижной оснастки ЛРК

В ЛРК заготовки деталей можно закреплять в специализированные подвижные оснастки. Это в свою очередь расширяет область доступа, а также значительно сокращает общее время обработки детали. В качестве подвижной оснастки в ЛРК выступает двухосевое позиционирующее устройство (рис. 4).

Рассмотрим задачу проектирования положения и ориентации в системе подвижной оснасткой. С точки зрения хранения данных в каждой точке технологической траектории появятся два дополнительных поля J_1 и J_2 , содержащих в себе значение поворота осей позиционера, но для генерации положения и ориентации инструмента необходимо провести пересчет.

Пусть векторы $\bar{p}'_0$ и $\bar{n}_0$ — положение и ориентация инструмента в точке технологической траектории, а j_1^0 и j_2^0 — нормированные векторы осей вращения позиционера в нейтральном положении (см. рис. 4).

В таком случае оси вращения j_1 и j_2 , соответствующие положению позиционера в углах J_1 и J_2 , будут иметь следующий вид:

$$j_1 = j_1^0; \quad (15)$$

$$j_2 = M(j_1^0, J_1). \quad (16)$$

Тогда для определения новых значений положения $\bar{p}'_n$ и ориентации $\bar{n}_n$ инстру-

мента необходимо осуществить поворот $\bar{p}'_0$ и $\bar{n}_0$ вокруг осей j_1 и j_2 .

Здесь же имеется особенность: поскольку при трансляции графических координат локального базиса модели в рабочую область ЛРК применяется метод переноса, описанный нами в статье [25], подразумевающий осуществление привязки модели и детали на основании трех пар точек пространства, постольку необходимо учесть значение углов J_1^0 и J_2^0 позиционера, при которых была осуществлена привязка.

В таком случае значения $\bar{p}'_n$ и $\bar{n}_n$ будут иметь следующий вид:

$$\bar{p}'_n = M(j_2, \pm(J_2 - J_2^0))M(j_1, \pm(J_1 - J_1^0))\bar{p}'_0; \quad (17)$$

$$\bar{n}_n = M(j_2, \pm(J_2 - J_2^0))M(j_1, \pm(J_1 - J_1^0))\bar{n}_0. \quad (18)$$

Причем в (17) и (18) знак $\pm$ выбирается в зависимости от кинематики движения позиционирующего устройства (в какую сторону осуществляется поворот устройства).

Стоит также отметить, что существуют подвижные оснастки, осуществляющие помимо вращательных движений, линейно-поступательные. В таком случае векторы $\bar{p}'_n$ и $\bar{n}_n$ в каждой точке технологической траектории будут сдвинуты на заданное значение линейно-поступательного смещения.

Положение инструмента относительно лазерного луча при проектировании технологической траектории

При проектировании ориентации инструмента в точках технологической траектории не учтён один факт, заключающийся в том, что одна степень свободы робота остаётся незафиксированной (вращение вокруг инструмента, которым в данном случае является лазерный луч). С точки зрения процесса лазерной обработки для неё можно было бы выбрать произвольное значение, т. к. поворот вокруг оси лазерного луча не влияет на сам процесс. Однако имеются технологические причины задавать для этого поворота определённое значение.

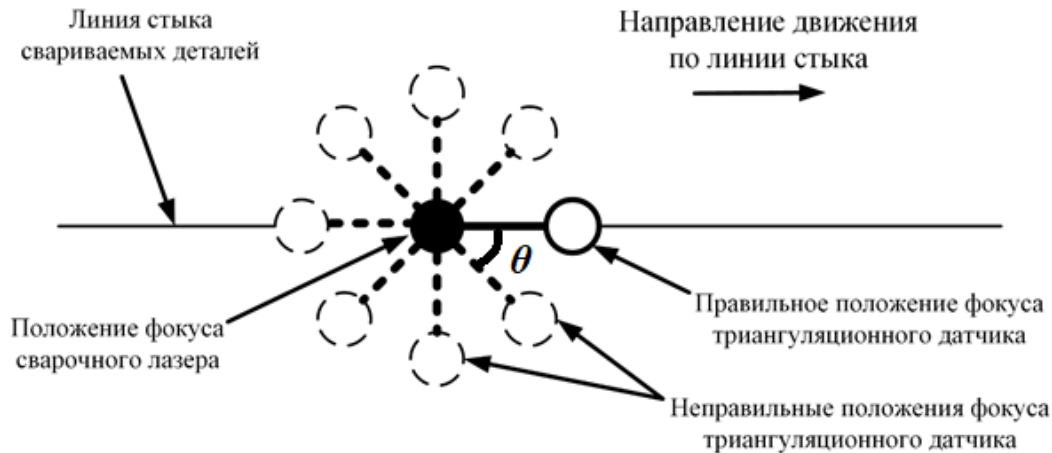


Рис. 7. Схема положения фокуса триангуляционного датчика при различном угле поворота последней оси робота (вид сверху)

Fig. 7. Diagram of the focus position of the triangulation sensor at different angles of rotation of the last axis of the robot (top view)

Изменение значения последней координаты будет приводить не только к вращению лазерной головки вокруг оси луча (что непринципиально), но и к вращению различного оборудования, смонтированного на инструменте. Пример такого оборудования – триангуляционный датчик, фокус которого вынесен в сторону от точки воздействия лазерного луча вокруг той же оси. Из этого следует, что в каждой точке технологической траектории угол поворота вокруг инструмента должен иметь определенное значение, задаваемое технологом (рис. 7).

В таком случае технолог должен задавать значение угла r в графическом пространстве САМ для каждой точки технологической траектории с учетом требуемого положения инструмента. Для автоматизации данного процесса расчет r возможно осуществить, основываясь на направлении движения инструмента.

Угол поворота r инструмента связан с направлением движения $\vec{d} = \{d_1, d_2, d_3\}$ следующим выражением:

$$\vec{d} = \{\sin(-r), \cos(-r), 0\}. \quad (19)$$

Из выражения (19) очевидным образом следует:

$$r = -\text{sign}(d_1) \cdot \arccos(d_2). \quad (20)$$

В таком случае положение угла r инструмента ЛРК вдоль направления $\vec{d}_i = p_i - p_{i-1}$

между смежными точками p_i и p_{i-1} технологической траектории с заданным отклонением на угол θ , примет следующий вид:

$$r = -\text{sign}(d_1^i \cos \theta - d_2^i \sin \theta) \times \arccos(d_1^i \sin \theta + d_2^i \cos \theta). \quad (21)$$

Выводы

Рассмотрен вопрос использования САМ-систем для промышленных лазерных роботизированных комплексов. Представлена наша программная разработка САМ-системы.

Изучена проблема построения точек технологической траектории в мировом графическом пространстве САМ относительно инструмента и подвижной оснастки ЛРК.

Для решения проблемы предложены методы расчета положения и ориентации инструмента лазерного роботизированного комплекса в локальном базисе модели.

Произведен расчет положения точек технологической траектории в системе подвижной оснастки ЛРК, построенных относительно инструмента.

Описан метод определения положения шестой оси промышленного робота (вращения инструмента вокруг лазерного луча) для задачи расчета позиционирования точки при построении технологической траектории относительно инструмента ЛРК в зависимости от направления движения вдоль зоны обработки.

Полученные результаты внедрены в САМ-систему ЛРК и успешно апробированы на лазерном роботизированном комплексе сварки. Для оценки качества полученного решения проведена серия из 16 испытаний, в ходе которых при различных положениях инструмента и позиционера производилось построение точек. Точки переносились в рабочее пространство ЛРК посредством подсистемы синхронизации модели и её детали. Результаты испытаний показали, что во всех случаях полученные значения координат полностью соответствовали спроектированным.

Результат построения точки технологической траектории относительно инструмента ЛРК в мировом графическом пространстве САМ проиллюстрирован на рис. 8. На нем изображена точка начала траектории А, успешно построенная в заданной позиции и соответствующей текущей ориентации инструмента.

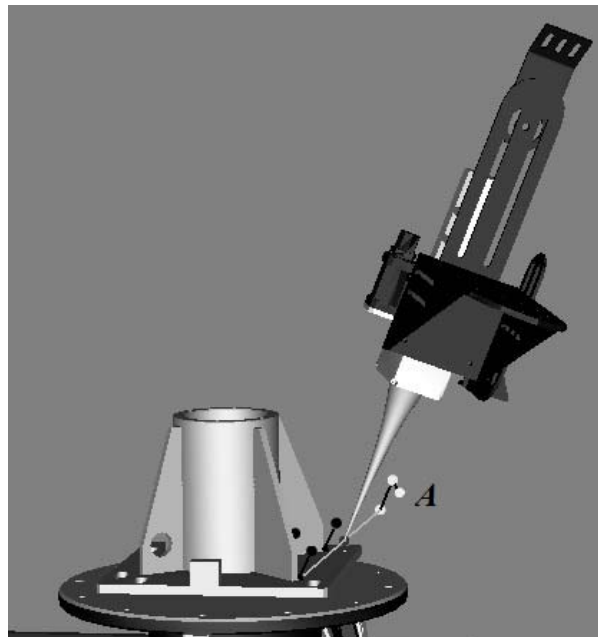


Рис. 8. Точка А, построенная на основании положения инструмента в мировом графическом пространстве САМ

Fig. 8. Point A, built on the basis of the position of the tool in the world graphic space CAM

СПИСОК ЛИТЕРАТУРЫ

1. Hui-Chi C., Guijun B., Chen-Nan S. High energy beam welding processes in manufacturing. Handbook of Manufacturing Engineering and Technology. 2014. Pp. 617–639. DOI: 10.1007/978-1-4471-4670-4_54
2. Игнатов А.Г. Российские лазерные технологии: состояние и перспективы применения // Ритм машиностроения. 2016. № 1. С. 26–40.
3. Lee H., Lim C.H.J., Low M.J., Tham N., Murukeshan V.M., Kim Y.J. Lasers in additive manufacturing: A review // Internat. J. of Precision Engineering and Manufacturing-Green Technology. 2017. Vol. 4. Iss. 3. Pp. 307–322. DOI: 10.1007/s40684-017-0037-7
4. Григорьев С.Н., Андреев А.Г., Ивановский С.П. Современное состояние и перспективы развития промышленной робототехники // Мехатроника, автоматизация, управление. 2013. № 1. С. 30–34
5. Кузнецов Д.А., Чернышев М.А., Овчинникова В.А. CAD САМ-системы для промышленных роботов // Интеллектуальный потенциал XXI века: ступени познания. 2017. № 35. С. 36–43
6. Авилов А.В., Каминская Ю.О., Трусова Д.С. Применение компьютерных систем для автоматизации и разработки новых технологий в машиностроительном производстве // Современные проблемы науки и образования. 2013. № 3. С. 16–22.
7. Zheng H., Cong M., Dong H., Liu Y., Liu D. CAD-based automatic path generation and optimization for laser cladding robot in additive manufacturing // The Internat. J. of Advanced Manufacturing Technology. 2017. Vol. 92. Iss. 9–12. Pp. 3605–3614. DOI: 10.1007/s00170-017-0384-0
8. Toquica J.S., Živanović S., Alvares A.J., Bonnard R. A STEP-NC compliant robotic machining platform for advanced manufacturing // The Internat. J. of Advanced Manufacturing Technology. 2017. Vol. 95. Iss. 9–12. Pp. 3839–3854. DOI: 10.1007/s00170-017-1466-8
9. Tonello S., Zanetti G.P., Finotto M., Bortolotto R., Tosello E., Menegatti E. WorkCellSimulator: A 3D simulator for intelligent manufacturing // Internat. Conf. on Simulation, Modeling, and Programming for Autonomous Robots. 2012. Pp. 311–322.
10. Ловыгин А.А. Будущее САМ-систем // Планета САМ 2018 // URL: http://planetacam.ru/articles/future_of_CAM.pdf (Дата обращения: 20.05.2019).

11. Iglesias I., Sebastián M.A., Aresc J.E. Overview of the state of robotic machining: Current situation and future potential // The Manufacturing Engineering Society Internat. Conf. MESIC. 2015. Pp. 911–917. DOI: 10.1016/j.proeng.2015.12.577
12. Устинова Л.Н. Особенности развития промышленности в условиях цифровизации // Формирование цифровой экономики и промышленности: новые вызовы. СПб.: Изд-во Политехнического ун-та, 2018. С. 176–197. DOI: 10.18720/IEP/2018.2/8
13. Bayram B., Ince G. Advances in robotics in the era of Industry 4.0. Industry 4.0: Managing the digital transformation. 2017. Pp. 187–200. DOI: 10.1007/978-3-319-57870-5_11
14. Mourtzis D., Makris S., Chryssolouris G. Computer-aided manufacturing // CIRP Encyclopedia of Production Engineering. 2018. DOI:10.1007/978-3-642-35950-7_6550-4
15. Lin W., Luo H. Robotic welding. Handbook of manufacturing engineering and technology. 2014. Pp. 1–36. DOI: 10.1007/978-1-4471-4976-7_106-1
16. Pan Z., Polden J., Larkin N., Duin S., Norrish J. Automated offline programming for robotic welding system with high degree of freedoms. Advances in computer, communication, control and automation. 2011. Pp. 685–692. DOI: 10.1007/978-3-642-25541-0_86
17. Milewski J. Computers, solid models, and robots // Additive manufacturing of metals. Springer Series in Materials Science. 2017. Vol. 258. Pp. 99–117. DOI: 10.1007/978-3-319-58205-4_6
18. Rodionov D., Lyukhter A., Prokoshev V. The mechanisms of constructing trajectories of a laser robotic complex for 3D polygonal models // J. of Physics: Conf. Series. 2018. DOI: 10.1088/1742-6596/1109/1/012007
19. Люхтер А.Б., Родионов Д.В., Лоханов А.В. Методы обработки информации, полученной посредством сканирования детали датчиком расстояния для оптимизации траектории // Перспективные технологии в средствах передачи информации. 2017. № 2. С. 138–141.
20. Лесков А.Г., Бажинова К.В., Морошкин С.Д., Феоктистова Е.В. Построение моделей кинематики исполнительных механизмов манипуляционных роботов с использованием блочных матриц // Инженерный журнал: наука и инновации. 2013. № 9. С. 58.
21. Булгаков А.Г., Воробьев В.Ф. Промышленные роботы. Кинематика, динамика, контроль и управление. М.: Солон-пресс, 2011. 486 с.
22. Лесков А.Г., Бажинова К.В., Селиверстова Е.В. Кинематика и динамика исполнительных механизмов манипуляционных роботов. М.: МГТУ им. Н.Э. Баумана, 2017. 107 с.
23. Фу К., Гонсалес Р., Ли К. Робототехника. Пер. с англ. М.: Мир, 1989. 624 с.
24. Flores P. Euler angles, Bryant angles and Euler parameters. Concepts and formulations for spatial multibody dynamics. 2015, Pp. 15–22. DOI: 10.1007/978-3-319-16190-7_4
25. Люхтер А.Б., Звягин М.Ю., Голубев А.С., Родионов Д.В. Перенос координат трехмерных CAD-моделей в пространство обрабатываемых деталей в роботизированных станочных комплексах // Динамика сложных систем XXI века. 2017. Т. 11. № 1. С. 40–46.
26. Rodionov D., Lyukhter A., Prokoshev V. 3D modeling of laser robotic complex motion in CAM spaces // Internat. Multi-Conf. on Industrial Engineering and Modern Technologies (FarEastCon), 2019, IEEE. DOI: 10.1109/FarEastCon.2019.8934248

Статья поступила в редакцию 20.05.2019.

REFERENCES

1. Hui-Chi C., Guijun B., Chen-Nan S. High energy beam welding processes in manufacturing. Handbook of manufacturing engineering and technology, 2014, Pp. 617–639. DOI: 10.1007/978-1-4471-4670-4_54
2. Ignatov A.G. Rossiyskiye lazernyye tekhnologii: sostoyaniye i perspektivy primeneniya [Russian laser technologies: state and prospects of application]. Ritm Mashinostroyeniya [Rhythm Engineering], 2016, No. 1, Pp. 26–40. (rus)
3. Lee H., Lim C.H.J., Low M.J., Tham N., Murukeshan V.M., Kim Y.J. Lasers in additive manufacturing: A review. International Journal of Precision Engineering and Manufacturing-Green Technology, 2017, Vol. 4, Issue 3, Pp. 307–322. DOI: 10.1007/s40684-017-0037-7
4. Grigoryev S.N., Andreyev A.G., Ivanovskiy S.P. Sovremennoye sostoyaniye i perspektivy razvitiya promyshlennoy robototekhniki [Current state and prospects of industrial robotics development]. Mekhatronika, Avtomatizatsiya, Upravleniye [Mechatronics, Automation, Control], 2013, No. 1, Pp. 30–34. (rus)
5. Kuznetsov D.A., Chernyshev M.A., Ovchinnikova V.A. CAD CAM sistemy dlya

promyshlennykh robotov [CAD CAM systems for industrial robots]. *Intellectualnyy Potentsial XXI Veka: Stupeni Poznaniya [Intellectual Potential of the XXI Century: Stages of Knowledge]*, 2017, No. 35, Pp. 36–43. (rus)

6. **Avilov A.V., Kaminskaya Yu.O., Trusova D.S.** Primeneniye kompyuternykh sistem dlya avtomatizatsii i razrabotki novykh tekhnologiy v mashinostroitel'nom proizvodstve [Application of computer systems for automation and development of new technologies in machine-building production]. *Sovremennyye Problemy Nauki i Obrazovaniya [Modern Problems of Science and Education]*, 2013, No. 3, Pp. 16–22. (rus)

7. **Zheng H., Cong M., Dong H., Liu Y., Liu D.** CAD-based automatic path generation and optimization for laser cladding robot in additive manufacturing. *The International Journal of Advanced Manufacturing Technology*, 2017, Vol. 92, Issue 9–12, Pp. 3605–3614. DOI: 10.1007/s00170-017-0384-0

8. **Toquica J.S., Živanović S., Alvares A.J., Bonnard R.** A STEP-NC compliant robotic machining platform for advanced manufacturing. *The International Journal of Advanced Manufacturing Technology*, 2017, Vol. 95, Issue 9–12, Pp. 3839–3854. DOI: 10.1007/s00170-017-1466-8

9. **Tonello S., Zanetti G.P., Finotto M., Bortoletto R., Tosello E., Menegatti E.** WorkCellSimulator: A 3D simulator for intelligent manufacturing. *International Conference on Simulation, Modeling, and Programming for Autonomous Robots*, 2012, Pp. 311–322

10. **Lovygin A.A.** Budushcheye CAM-sistem. Planeta CAM 2018. Available: http://planetacam.ru/articles/future_of_CAM.pdf (Accessed: 20.05.2019). (rus)

11. **Iglesias I., Sebastián M.A., Aresc J.E.** Overview of the state of robotic machining: Current situation and future potential. *The Manufacturing Engineering Society International Conference, 2015*, Pp. 911–917. DOI: 10.1016/j.proeng.2015.12.577

12. **Ustinova L.N.** Osobennosti razvitiya promyshlennosti v usloviyakh tsifrovizatsii [Features of industry development in the conditions of digitalization]. *Formirovaniye Tsifrovoy Ekonomiki i Promyshlennosti: Novyye Vyzovy [Formation of Digital Economy and Industry: New Challenges]*. St. Petersburg: Polytechnic University Publ., 2018, Pp. 176–197. (rus) DOI: 10.18720/IEP/2018.2/8

13. **Bayram B., Ince G.** *Advances in robotics in the Era of Industry 4.0. Industry 4.0: Managing the digital transformation*. 2017, Pp. 187–200. DOI: 10.1007/978-3-319-57870-5_11

14. **Mourtzis D., Makris S., Chrysosouris G.** Computer-aided manufacturing. *CIRP Encyclopedia of Production Engineering*. 2018. DOI:10.1007/978-3-642-35950-7_6550-4

15. **Lin W., Luo H.** *Robotic welding. Handbook of manufacturing engineering and technology*. 2014, Pp. 1–36. DOI: 10.1007/978-1-4471-4976-7_106-1

16. **Pan Z., Polden J., Larkin N., Duin S., Norrish J.** *Automated offline programming for robotic welding system with high degree of freedoms. Advances in computer, communication, control and automation*. 2011, Pp. 685–692. DOI: 10.1007/978-3-642-25541-0_86

17. **Milewski J.** Computers, solid models, and robots. *Additive Manufacturing of Metals. Springer Series in Materials Science*, 2017, Vol. 258, Pp. 99–117. DOI: 10.1007/978-3-319-58205-4_6

18. **Rodionov D., Lyukhter A., Prokoshev V.** The mechanisms of constructing trajectories of a laser robotic complex for 3D polygonal models. *Journal of Physics: Conf. Series*, 2018. DOI: 10.1088/1742-6596/1109/1/012007

19. **Lyukhter A.B., Rodionov D.V., Lokhanov A.V.** Metody obrabotki informatsii, poluchennoy posredstvom skanirovaniya detali datchikom rassoyaniya dlya optimizatsii trayektorii [Methods of processing information obtained by scanning a part with a distance sensor to optimize the trajectory]. *Perspektivnyye Tekhnologii v Sredstvakh Peredachi Informatsii [Promising Technologies in Information Transmission Media]*, 2017, No. 2, Pp. 138–141. (rus)

20. **Leskov A.G., Bazhinova K.V., Moroshkin S.D., Feoktistova Ye.V.** Postroyeniye modeley kinematiki ispolnitelnykh mekhanizmov manipulyatsionnykh robotov s ispolzovaniyem blochnykh matrits [Construction of models of kinematics of executive mechanisms of manipulative robots using block matrices]. *Inzhenernyy Zhurnal: Nauka i Innovatsii [Engineering Journal: Science and Innovation]*, 2013, No. 9, P. 58. (rus)

21. **Bulgakov A.G., Vorobyev V.F.** *Promyshlennyye roboty. Kinematika, dinamika, kontrol i upravleniye [Industrial robot. Kinematics, dynamics, control and management]*. Moscow: Solon-press, 2011, 486 p. (rus)

22. **Leskov A.G., Bazhinova K.V., Seliverstova Ye.V.** *Kinematika i dinamika ispolnitelnykh mekhanizmov manipulyatsionnykh robotov [Kinematics and dynamics of executive mechanisms of manipulation robots]*. Moscow: MGTU Publ., 2017, 107 p. (rus)

23. **Fu K., Gonsales R., Li K.** *Robototekhnika [Radioengineering]*. Moscow: Mir Publ., 1989, 624 p. (rus)

24. Flores P. *Euler angles, Bryant angles and Euler parameters. Concepts and formulations for spatial multibody dynamics*. 2015, Pp. 15–22. DOI: 10.1007/978-3-319-16190-7_4

25. Lyukhter A.B., Zvyagin M.Yu., Golubev A.S., Rodionov D.V. Perenos koordinat trekhmernykh CAD-modeley v prostranstvo obrabatyvayemykh detaley v robotizirovannykh stanochnykh kompleksakh [The transfer of coordinates of three-dimensional CAD model in the space of a

workpiece in a machining robotic complex]. *Dinamika Slozhnykh Sistem XXI Veka [Dynamics of Complex Systems-XXI Century]*, 2017, Vol. 11, No. 1, Pp. 40–46. (rus)

26. Rodionov D., Lyukhter A., Prokoshev V. 3D modeling of laser robotic complex motion in CAM spaces. *2019 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*, 2019, IEEE. DOI: 10.1109/FarEastCon.2019.8934248

Received 20.05.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

РОДИОНОВ Дмитрий Викторович

RODIONOV Dmitrii V.

E-mail: d3391rod@gmail.com

ЛЮХТЕР Александр Борисович

LYUKHTER Alexander B.

E-mail: 3699137@mail.ru

ПРОКОШЕВ Валерий Григорьевич

PROKOSHEV Valery G.

E-mail: prokoshev_vg@vlsu.ru

DOI: 10.18721/JCSTCS.12411
УДК 681.51.11

СИНТЕЗ ДИСКРЕТНОГО РЕГУЛЯТОРА ПЕРЕОБОРУДОВАНИЕМ С УЧЁТОМ ЗАДЕРЖКИ В ВЫЧИСЛИТЕЛЬНОМ УСТРОЙСТВЕ

В.М. Филиповский¹, Н.Л. Полюга², Б.Б. Мухамбедьяров²

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация;

² Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики,
Санкт-Петербург, Российская Федерация

Рассмотрена методика синтеза дискретного регулятора электромеханической системы методами переоборудования: Эйлера, обратных разностей, Тастина. Проведен сравнительный анализ полученных дискретных систем. На основе результатов экспериментов, выполненных в среде Simulink, показана необходимость учёта времени запаздывания в цифровом вычислительном устройстве еще на этапе синтеза непрерывного регулятора. Представлены сравнительные зависимости показателей качества синтезированной системы от интервала квантования T дискретной части системы автоматического управления при различных методах переоборудования. Показано, что с увеличением интервала квантования в дискретной системе возрастает перерегулирование. С ростом T наблюдается увеличение запаздывания реакции дискретной системы на единичную ступенчатую функцию по отношению к реакции непрерывной системы. Наилучшие показатели получены при переоборудовании регулятора методом обратных разностей, наихудшие — методом Эйлера.

Ключевые слова: электромеханическая система, двигатель постоянного тока, переоборудование, дискретный регулятор, задержка, показатели качества.

Ссылка при цитировании: Филиповский В.М., Полюга Н.Л., Мухамбедьяров Б.Б. Синтез дискретного регулятора переоборудованием с учётом задержки в вычислительном устройстве // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 136–144. DOI: 10.18721/JCSTCS.12411

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

SYNTHESIS OF THE DISCRETE REGULATOR BY RE-EQUIPMENT METHOD TAKING INTO ACCOUNT THE DELAY IN A COMPUTER DEVICE

V.M. Filipovskii¹, N.L. Polyuga², B.B. Mukhambedyarov²

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation;

² St. Petersburg National Research University of Information Technologies,
Mechanics and Optics, St. Petersburg, Russian Federation

The technique of synthesis of a discrete regulator of an electromechanical system by the methods of re-equipment is considered: Euler, inverse differences, Tustin. A comparative analysis of the discrete systems obtained was carried out. Based on the results of experiments performed in the Simulink environment, the necessity of taking into account the lag time in a digital computing device at the stage of synthesis of a

continuous regulator is shown. The comparative dependences of the quality indicators of the synthesized system on the quantization interval T of the discrete part of the automatic control system with different methods of conversion are presented. It is shown that with an increase in the quantization interval in a discrete system, the overshoot increases. With the growth of T , there is an increase in the delay in the reaction of a discrete system to a single step function with respect to the reaction of a continuous system. The best results are obtained by the conversion of the regulator using the method of inverse differences, the worst – the Euler method.

Keywords: electromechanical system, DC motor, re-equipment, discrete controller, delay, quality indicators.

Citation: Filipovskii V.M., Polyuga N.L., Mukhambedyarov B.B. Synthesis of the discrete regulator by re-equipment method taking into account the delay in a computer device. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 136–144. DOI: 10.18721/JCSTCS.12411

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

При построении систем автоматического управления важную роль играют корректирующие устройства (регуляторы), необходимые для придания системам желаемых статических и динамических характеристик. В современных системах управления вместо аналоговых применяются цифровые регуляторы, реализованные с помощью микропроцессоров, микроконтроллеров, бортовых компьютеров, программируемых логических контроллеров, благодаря их преимуществам [1]:

- отсутствует дрейф параметров элементов регулятора;
- имеется возможность реализации сложных алгоритмов с очень высокими массогабаритными показателями;
- для изменения алгоритма управления достаточно изменить программу и загрузить её в цифровое устройство.

Однако любая дискретная система имеет важный параметр – интервал квантования T , с которым связана другая особенность цифровой системы: задержка между моментом поступления сигнала на аналогово-цифровой преобразователь вычислительного устройства и появлением управляющего воздействия на его выходе вследствие конечного времени вычисления данного воздействия [2]. Отдельные вопросы, связанные с временной задержкой в дискретной системе, рассмотрены в [2–8].

В работах [6, 9] показаны методики синтеза цифровой системы заменой непрерывного регулятора дискретным эквивалентом. По принятой терминологии [9, 10] данный метод называется *переоборудованием*. При этом не рассмотрен синтез цифрового регулятора с учётом временного запаздывания в вычислительном устройстве, не проведен сравнительный анализ синтезированных систем при различных способах переоборудования.

В настоящей работе предложена методика учёта временной задержки при синтезе дискретного регулятора электромеханической системы. Также исследовано влияние величины интервала квантования в вычислительном устройстве на качество дискретных систем при разных методах переоборудования: Эйлера, обратных разностей, Тастина.

Постановка задачи синтеза дискретных систем управления

Математическая модель непрерывной электромеханической системы. В качестве объекта управления рассматривается электрический двигатель постоянного тока с силовым преобразователем [4–6].

Объект управления, в соответствии со структурной схемой (рис. 1), описывается передаточной функцией (ПФ):

$$W_o(s) = \frac{k_{\Pi} k_M}{(J R_{\text{я}} T_{\text{я}} s^2 + J R_{\text{я}} s + k_E k_M)(1 + T_{\Pi} s)}. \quad (1)$$

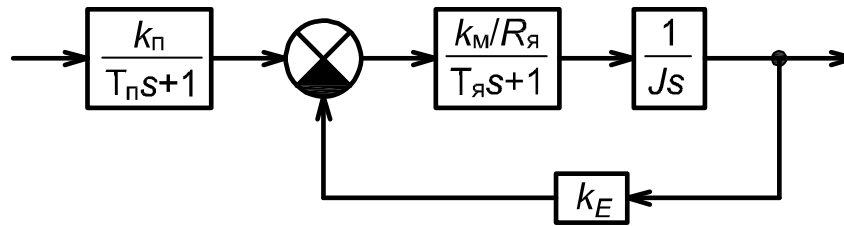


Рис. 1. Структурная схема объекта управления

Fig. 1. Structural scheme of control object

Если обозначить:

$$T_M = \frac{J R_Я}{k_E k_M}, \quad (2)$$

то выражение (1) можно записать в виде:

$$W_0(s) = \frac{k_П / k_E}{(T_M T_Я s^2 + T_M s + 1)(T_П s + 1)}, \quad (3)$$

или

$$W_0(s) = \frac{k_П / k_E}{(1 + T_1 s)(1 + T_2 s)(T_П s + 1)}, \quad (4)$$

которое с учётом численных значений параметров системы (рис. 1):

$$k_П = 20, J = 0,05, R_Я = 0,01 \text{ Гн}, \\ k_E = 1, k_M = 1, R_Я = 1,$$

$$T_Я = \frac{L_Я}{L_Я} = 0,01 \text{ с}, T_П = \frac{1}{150} = 0,0067 \text{ с}, T_M = 0,05 \text{ с}$$

может быть записано как

$$W_0(s) = \frac{20}{(1 + 0,0362s)(1 + 0,0138s)(1 + 0,0067s)}. \quad (5)$$

Синтез цифровой системы методом переоборудования

Рассматриваемый метод синтеза дискретного регулятора, именуемый переоборудованием, предполагает преобразование регулятора, синтезированного для непрерывной системы, в дискретную форму.

Настройка непрерывной системы на модульный оптимум. Критерием синтеза регулятора для непрерывной системы выбрана настройка последней на модульный оптимум (МО).

Желаемая передаточная функция разомкнутой системы, настроенной на МО, должна иметь вид [12]:

$$W_{MO}(s) = \frac{1}{2T_\mu s(1 + T_\mu s)}, \quad (6)$$

где T_μ – сумма малых постоянных времени системы управления.

При этом передаточная функция регулятора при выбранной настройке определяется как

$$W_P(s) = \frac{k_E(1 + T_1 s)}{2k_П T_\mu s}. \quad (7)$$

При определении ПФ регулятора постоянные времени объекта управления (см. выражение (5)) были разделены на большие и малые. Большой выбрана постоянная T_1 , а к малым отнесены T_2 и $T_П$, которые в синтезированной системе представлены эквивалентной постоянной $T_\mu = T_2 + T_П$; $T_\mu = 0,0205 \text{ с}$.

С целью проверки правильности настройки выполнено моделирование синтезированной системы. Переходная характеристика, изображенная на рис. 2, определяется следующими показателями: перерегулирование σ составляет 4,6 %, время переходного процесса $t_{пп}$ равняется 0,171 с = 8,34 T_μ , а время первого согласования ($t_{пс}$) 0,089 с = 4,34 T_μ . Показатели качества замкнутой системы незначительно отличаются от теоретических ($\sigma = 4,3 \%$, $t_{пп} = 8,4T_\mu$, $t_{пс} = 4,7T_\mu$), что свидетельствует о правильной настройке системы.

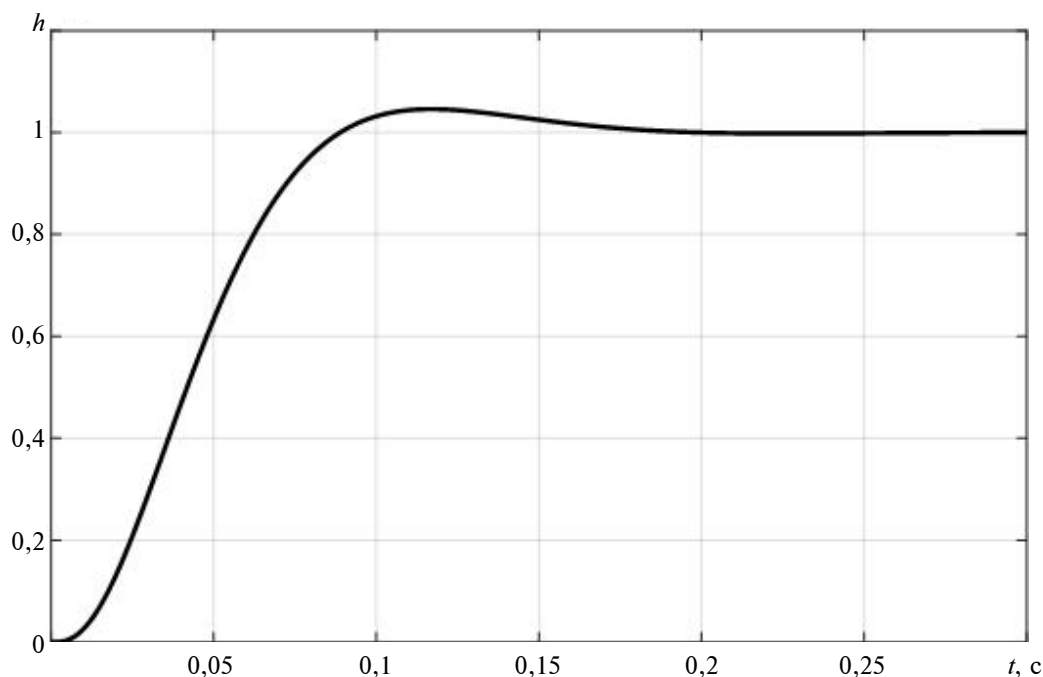


Рис. 2. Переходная характеристика синтезированной системы

Fig. 2. Transient characteristic of the synthesized system

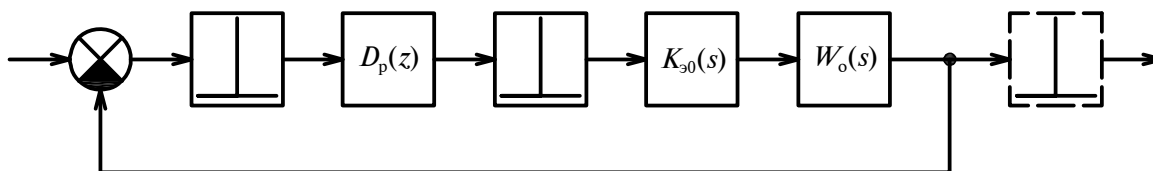


Рис. 3. Структурная схема дискретной системы автоматического управления

Fig. 3. Structural scheme of the automation control discrete system

В реальной дискретной системе регулятор обычно реализуется на базе цифрового бортового компьютера или микроконтроллера, в котором имеет место задержка между моментом поступления сигнала на аналогово-цифровой преобразователь вычислительного устройства и появлением управляющего воздействия на его выходе вследствие конечного времени вычисления данного воздействия, соответствующего интервалу квантования T дискретной системы.

Следовательно, эту задержку необходимо заранее учесть в передаточной функции объекта, при синтезе непрерывного регулятора. В объекте управления с ПФ (1)–(5) её можно в простейшем случае представить

дополнительным аperiодическим звеном с постоянной времени, равной T , которая увеличивает значение суммы малых постоянных T_μ . С учетом сказанного выше ПФ регулятора примет вид:

$$W_p(s) = \frac{k_E (1 + T_I s)}{2k_\Pi (T_\mu + T)s}. \quad (8)$$

Пересчет математической модели непрерывного регулятора к дискретному виду. Математическая модель дискретной системы (рис. 3) будет иметь ту же структуру, что и непрерывная модель замкнутой электромеханической системы.

Отличием является дискретный регулятор (ДР) вместо непрерывного, присут-

ствие двух квантователей на входе и выходе цифрового регулятора, экстраполятора нулевого порядка и квантователя на выходе непрерывного объекта.

ДР будет получен из (8) методами переоборудования. Задача переоборудования состоит в том, чтобы заменить спроектированный непрерывный регулятор цифровым устройством так, чтобы сохранить все существенные свойства непрерывной системы (устойчивость, точность, быстродействие).

Методы переоборудования основаны на приближённой замене интегрирующего звена с передаточной функцией $1/s$ его дискретной моделью. В зависимости от способа вычисления определённого интеграла получаются соответствующие методы переоборудования.

Правило для дискретизации аналогового регулятора методом переоборудования в общем случае имеет вид:

$$D_p(z) = W_p(s)|_{s=I(z)},$$

где $I(z)$ — выражение, на которое заменяется s в ПФ непрерывного регулятора.

В работе использованы следующие методы переоборудования: метод Эйлера, метод обратных разностей, метод Тастина.

Метод Эйлера заключается в том, что в ПФ регулятора s необходимо заменить на выражение $s \rightarrow \frac{z-1}{T}$.

При переоборудовании методом обратных разностей необходимо воспользоваться заменой $s \rightarrow \frac{z-1}{zT}$.

При синтезе методом Тастина $s \rightarrow \frac{2z-1}{Tz+1}$.

После переоборудования приведёнными методами соответственно получаются передаточные функции дискретного регулятора:

$$D_p(z) = \frac{k_E(T_1 z + T - T_1)}{2k_{\Pi}(T_{\mu} + T)(z-1)}, \quad (9)$$

$$D_p(z) = \frac{k_E((T_1 + T)z - T_1)}{2k_{\Pi}(T_{\mu} + T)(z-1)}, \quad (10)$$

$$D_p(z) = \frac{k_E((T + 2T_1)z + T - 2T_1)}{4k_{\Pi}(T + T_{\mu})(z-1)}. \quad (11)$$

Передаточные функции дискретных регуляторов без учёта задержки могут быть получены обнулением T в знаменателе соответствующей передаточной функции дискретного регулятора (9)–(11).

Исследование синтезированной цифровой электромеханической системы

На первом этапе исследования сравниваются показатели качества системы по переходным характеристикам — перерегулирования, времени переходного процесса, времени первого согласования аналоговой и дискретных систем с ПФ дискретных регуляторов (9)–(11) при следующих значениях интервала квантования: $T = 0,00105$ с (соответствует $0,05T_{\mu}$) (рис. 4 а); $T = T_{\mu}$ (рис. 4 б).

На рисунке показано, что с увеличением T наблюдается ухудшение динамики дискретной системы, отчётливо видна задержка в отработке входного сигнала. При $T = T_{\mu}$ переходные характеристики имеют следующие показатели:

метод Эйлера: $\sigma = 17,9 \%$, $t_{\text{nc}} = 0,149$ с, $t_{\text{nn}} = 0,43$ с;

метод обратных разностей: $\sigma = 8,5 \%$, $t_{\text{nc}} = 0,129$ с, $t_{\text{nn}} = 0,31$ с;

метод Тастина: $\sigma = 12,3 \%$, $t_{\text{nc}} = 0,139$ с, $t_{\text{nn}} = 0,39$ с.

Далее для сравнения показателей качества проведено моделирование аналоговой и дискретных систем с ПФ дискретных регуляторов, синтезированных без учёта задержки, при $T = 0,05T_{\mu}$ (рис. 5 а) и $T = T_{\mu}$ (рис. 5 б).

При интервале квантования равном $0,05T_{\mu}$ все кривые практически совпадают, как и на рис. 4 а. Однако при $T = T_{\mu}$ кривые существенно отличаются от изображённых на рис. 4 б и имеют следующие показатели:

метод Эйлера: $\sigma = 62,4 \%$, $t_{\text{nc}} = 0,09$ с, $t_{\text{nn}} = 0,94$ с;

метод обратных разностей: $\sigma = 61 \%$, $t_{\text{nc}} = 0,075$ с, $t_{\text{nn}} = 0,93$ с;

метод Тастина: $\sigma = 59,2 \%$, $t_{\text{nc}} = 0,129$ с, $t_{\text{nn}} = 0,92$ с.

Дискретная система с регулятором, рассчитанным без учета времени задержки, при относительно больших T становится существенно колебательной по сравнению с аналоговой системой.

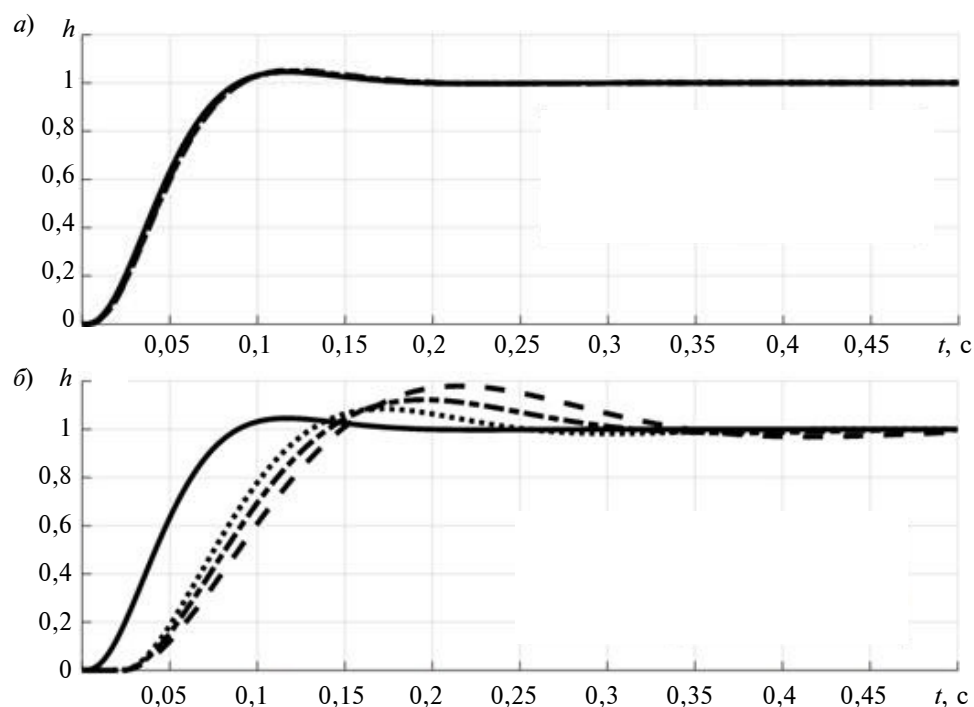


Рис. 4. Переходные характеристики дискретных САУ с учётом задержки в вычислительном устройстве
(—) — аналоговая система; (---) — метод Эйлера; (·····) — метод обратных разностей;
(-·-·-) — метод Тастина

Fig. 4. Transient characteristics of automation control discrete systems taking into delay in computing device

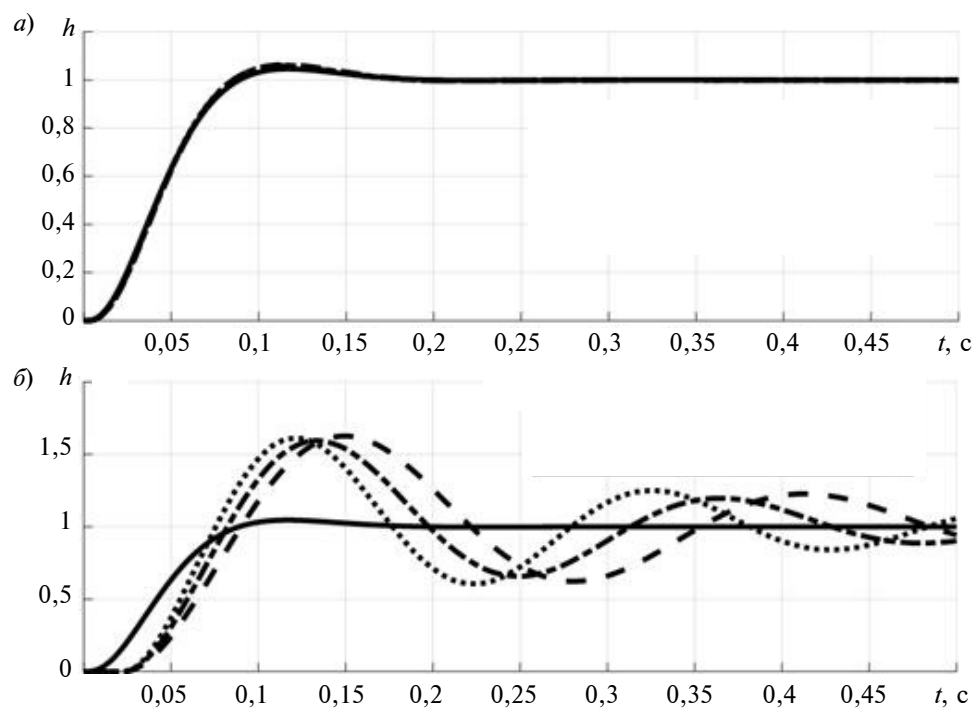


Рис. 5. Переходные характеристики дискретных САУ без учёта задержки в вычислительном устройстве
(—) — аналоговая система; (---) — метод Эйлера; (·····) — метод обратных разностей;
(-·-·-) — метод Тастина

Fig. 5. Transient characteristics of automation control discrete systems without taking into delay in computing device

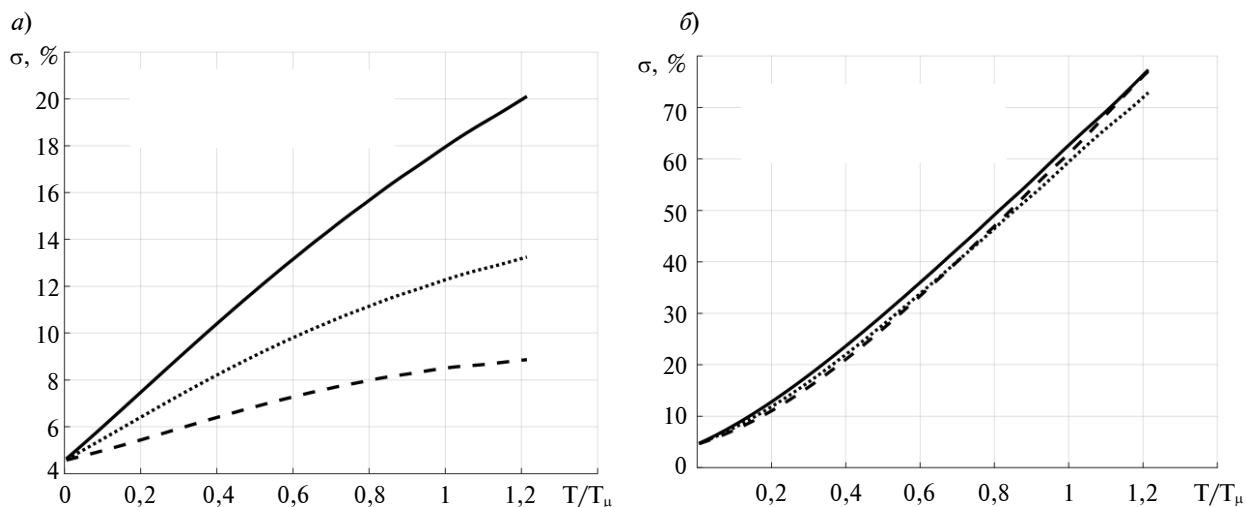


Рис. 6. Зависимости перерегулирования от интервала квантования в дискретных САУ

(—) — метод Эйлера; (---) — метод обратных разностей; (.....) — метод Тастина

Fig. 6. Dependences of overshoot from quantization interval in automation control discrete systems

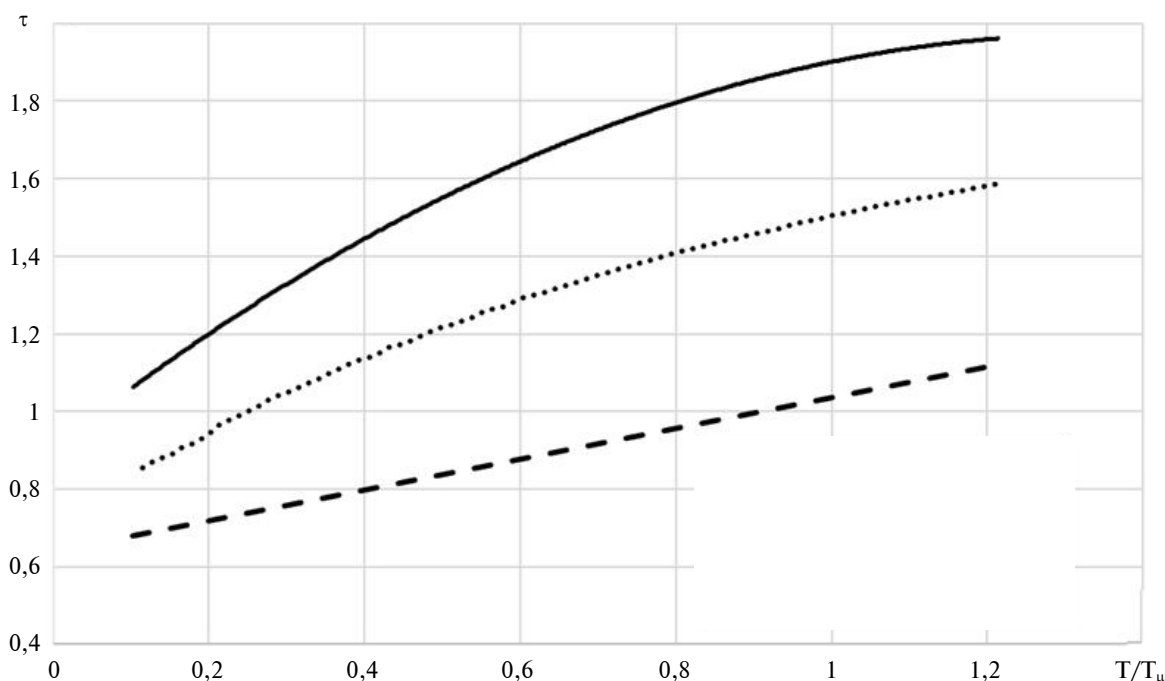


Рис. 7. Запозывание реакции дискретной системы

(—) — метод Эйлера; (---) — метод обратных разностей; (.....) — метод Тастина

Fig. 7. Delay of response of discrete system

Следующий этап исследования представляет собой более подробное изучение влияния величины T на показатели качества у системы с передаточными функциями регуляторов (9)–(11). При моделировании интервал квантования в цифровой

системе изменялся от $T = 0,005T_\mu$ до $1,2T_\mu$. Показаны графики зависимости σ от T/T_μ (рис. 6 а). Для сравнения представлены (рис. 6 б) графики для систем, полученных без учёта времени запаздывания при синтезе регулятора.

На рис. 6 б при увеличении T наблюдаются большие σ , что недопустимо. Причём данные показатели, полученные разными методами, отличаются незначительно по сравнению с величинами на рис. 6 а.

Также представлены (рис. 7) кривые зависимости относительного запаздывания τ переходной характеристики дискретной системы ($t_{пс.дс}$) по отношению к непрерывной ($t_{пс.ас}$). Этот показатель определяется выражением:

$$\tau = \frac{t_{пс.дс} - t_{пс.ас}}{T}.$$

Видно, что запаздывание реакции дискретной системы возрастает с увеличением интервала квантования. Наибольшее запаздывание наблюдается у системы с регулятором, полученным методом Эйлера, наименьшее — методом обратных разностей.

Выводы

Из полученных результатов следует, что если интервал квантования $T \ll T_u$, то показатели качества дискретной системы практически совпадают с показателями аналоговой независимо от метода переоборудования и учёта задержки при синтезе цифрового регулятора.

С увеличением T наблюдается ухудшение качества дискретной системы по отношению к аналоговой. В системе, где регулятор синтезирован с учётом времени задержки, увеличение значений показателей качества пе-

реходной характеристики (перерегулирования, времени первого согласования) относительно показателей непрерывной системы при изменении интервала квантования носит практически линейный характер. При $T = T_u$ перерегулирование возрастает до 18 % (метод Эйлера). Наилучший показатель у системы, переоборудованной методом обратных разностей ($\sigma \approx 8,5\%$). Имеет место запаздывание реакции дискретной системы на единичную ступенчатую функцию по отношению к реакции непрерывной системы. Причём данное запаздывание наибольшее для системы, где регулятор получен методом Эйлера (достигает $2T$), и наименьшее, когда переоборудование выполнено методом обратных разностей ($\approx T$).

Если провести синтез регулятора без принятия во внимание временной задержки вычислительного устройства, то в дискретной системе с ростом T наблюдается существенное ухудшение показателей качества по переходным характеристикам по сравнению с рассматриваемой в работе системой (когда задержка в регуляторе учтена): значительно возрастают перерегулирование и время переходного процесса, система становится более колебательной.

Таким образом, система с регулятором, полученным методом обратных разностей, имеет наилучшие показатели качества. Время запаздывания в вычислительном устройстве крайне необходимо учитывать при синтезе цифрового регулятора.

СПИСОК ЛИТЕРАТУРЫ

1. Олссон Г., Пиани Д. Цифровые системы автоматизации и управления. СПб.: Невский диалект, 2001. 557 с.
2. Gielen R.H. Stability analysis and control of discrete-time systems with delay. Eindhoven: Technische Universiteit Eindhoven, 2013. 169 p. DOI: 10.6100/IR747896.
3. Haugen F. Discrete-time signals and systems. Skien: TechTeach, 2005. 75 p.
4. Bedoui S., Ltaief M., Abderrahim K. New results on discrete-time delay systems identification // Internat. J. of Automation and Computing. 2012. No.9(6). Pp. 570–577. DOI: 10.1007/s11633-012-0681-x.
5. Debeljkovic D.L., Stojanovic S.B. The stability of linear discrete time delay systems in the sense of Lyapunov: An overview // Scientific Technical Review. 2010. Vol. 60. No. 3–4. Pp. 67–81.
6. H'Mida B., Sahbi M., Dhaou S. Discretizing of linear systems with time-delay using method of Euler's and Tustin's approximations // Internat. J. of Engineering Research and Applications. 2015. Vol. 5. No. 3. Pp. 83–89.
7. Stojanovic S.B., Debeljkovic D.L., Dimitrijevic N. Stability of discrete-time systems with time-varying delay: Delay decomposition approach // Internat. J. of Computers, Communications and Control. 2012. Vol. 7. No. 4. Pp. 776–784.
8. Silva L.F.P., Leite V.J.S., Castelan E.B., Lopes A.N.D. Stability and controller design for T-S fuzzy discrete-time systems with time-varying delay in the state // 2018 IEEE Internat. Conf. on Fuzzy Systems. Rio de Janeiro, 2018. Pp. 1–7. DOI: 10.1109/FUZZ-IEEE.2018.8491660.
9. Бобиков А.И., Борисов М.М. Построение цифровых регуляторов путем переоборудования непрерывных регуляторов // Сб. тр. II Междунар. науч.-техн. и науч.-метод. конф.: в 8 т. Т. 4. Рязань, 2018. С. 58–62.

10. Поляков К.Ю. Полиномиальный синтез оптимальных цифровых следящих систем. I // АИТ. 2001. № 2. С. 149–182.

11. Полюга Н.Л., Филиповский В.М. Цифровой регулятор в системах автоматического управления // Информатика и кибернетика: сб. докладов студенческой науч. конф. Института

компьютерных наук и технологий. СПб., 2016. С. 133–137.

12. Кояин Н.В., Мальцева О.П., Удут Л.С. Оптимизация контуров регулирования систем электропривода по типовым методикам // Известия Томского политехнического университета. 2005. Т. 308. № 7. С. 120–125.

Статья поступила в редакцию 06.05.2019.

REFERENCES

1. Olsson G., Piani D. *Tsifrovyye sistemy avtomatizatsii i upravleniya* [Digital automation and control systems]. St. Petersburg: Nevskiy Dialect Publ., 2001, 557 p. (rus)

2. Gielen R.H. *Stability analysis and control of discrete-time systems with delay*. Eindhoven: Technische Universiteit Eindhoven, 2013, 169 p. DOI: 10.6100/IR747896.

3. Haugen F. *Discrete-time signals and systems*. Skien: TechTeach, 2005. 75 p.

4. Bedoui S., Ltaief M., Abderrahim K. New results on discrete-time delay systems identification. *International Journal of Automation and Computing*, 2012, No. 9 (6), Pp. 570–577. DOI: 10.1007/s11633-012-0681-x.

5. Debeljkovic D.L., Stojanovic S.B. The stability of linear discrete time delay systems in the sense of Lyapunov: An overview. *Scientific Technical Review*, 2010, Vol. 60, No. 3–4, Pp. 67–81.

6. H'Mida B., Sahbi M., Dhaou S. Discretizing of linear systems with time-delay using method of Euler's and Tustin's approximations. *International Journal of Engineering Research and Applications*, 2015, Vol. 5, No. 3, Pp. 83–89.

7. Stojanovic S.B., Debeljkovic D.L., Dimitrijevic N. Stability of discrete-time systems with time-varying delay: Delay decomposition approach. *International Journal of Computers, Communications and Control*, 2012, Vol. 7, No. 4, Pp. 776–784.

8. Silva L.F.P., Leite V.J.S., Castelan E.B., Lopes A.N.D. Stability and controller design for T-S fuzzy discrete-time systems with time-varying delay in the state. *2018 IEEE International Conference on Fuzzy Systems*, 8–13 July 2018, Rio de Janeiro, Pp. 1–7. DOI: 10.1109/FUZZ-IEEE.2018.8491660.

9. Bobikov A.I., Borisov M.M. Postroyeniye tsifrovyykh regulyatorov putem pereoborudovaniya nepreryvnykh regulyatorov [Construction of digital controllers by conversion of continuous controllers]. *Proceedings of 2nd International Scientific Technical and Scientific Methodical Conference on Sovremennyye tekhnologii v nauke i obrazovanii*. Ryazan, 2018, Vol. 4, Pp. 58–62. (rus)

10. Polyakov K.Yu. Polinomialnyy sintez optimalnykh tsifrovyykh sledyashchikh sistem. I. [Polynomial Design of Optimal Sampled-Data Tracking Systems. I]. *Automation and Remote Control*, 2001, No. 2, Pp. 149–182. (rus)

11. Polyuga N.L., Filipovskiy V.M. Tsifrovoy regulyator v sistemakh avtomaticheskogo upravleniya [Digital controller in automatic control systems]. *Informatika i kibernetika. ComCon-2016*, St. Petersburg, 3–8 April 2016, Pp. 133–137. (rus)

12. Koyain N.V., Maltseva O.P., Udut L.S. Optimizatsiya konturov regulirovaniya sistem elektroprivoda po tipovym metodikam [Optimization of control circuits of electric drive systems by typical methods]. *Izvestiya Tomskogo Politehnicheskogo Universiteta*. 2005, Vol. 308, No. 7, Pp. 120–125. (rus)

Received 06.05.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

ФИЛИПОВСКИЙ Владимир Михайлович

FILIPOVSKIY Vladimir M.

E-mail: fil@icc.spbstu.ru

ПОЛЮГА Николай Любомирович

POLYUGA Nikolay L.

E-mail: van-der-gan@mail.ru

МУХАМБЕДЬЯРОВ Бекбол Бокейханович

MUKHAMBEDYAROV Bekbol B.

E-mail: space_line93@mail.ru



DOI: 10.18721/JCSTCS.12412
УДК 004.021

ОБЗОР ПОДХОДОВ К ОБНАРУЖЕНИЮ СБОЕВ В СИСТЕМАХ ХРАНЕНИЯ ДАННЫХ

М.Б. Успенский

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Выполнен обзор существующих программных средств, предназначенных для мониторинга состояния систем хранения данных, определены применяющиеся подходы к сбору, обработке и хранению данных, описаны используемые инструменты для обнаружения сбоев, сформулированы перечни признаков для классификации и сравнения существующих программных решений. На основании проведенного анализа решаемых существующими программными средствами задач предложена типовая архитектура программного комплекса для обнаружения сбоев, описаны входящие в неё модули и характер их взаимодействия. Выполнен обзор актуальных публикаций, посвященных обнаружению сбоев и выявлению аномалий в сфере хранения данных и вычислительных систем, рассмотрены представленные в них алгоритмы, основанные на методах кластеризации и классификации, статистического анализа, опорных векторов, изолирующего леса, искусственных иммунных систем, сетей инвариантов и др.

Ключевые слова: выявление аномалий, машинное обучение, обнаружение сбоев, система хранения данных, локализация неисправностей.

Ссылка при цитировании: Успенский М.Б. Обзор подходов к обнаружению сбоев в системах хранения данных // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2019. Т. 12. № 4. С. 145–158. DOI: 10.18721/JCSTCS.12412

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>)

A SURVEY OF THE APPROACHES TO STORAGE SYSTEMS FAULT DETECTION

M.B. Uspenskiy

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

In present paper, we have carried out a comparative analysis of existing software used for health monitoring in enterprise-level storage systems, described commonly used approaches to monitoring data collection, processing and storage, fault detection methods. Based on this analysis we proposed criteria for monitoring software classification and comparison, generalized monitoring software architecture, its modules and module interaction. We also carried out a survey of the recent publications dedicated to anomalies detection, fault diagnosis in a field of data storage and computing systems, and described commonly used algorithms, including clusterization and classification methods, statistical analysis, SVM, isolated forest, artificial immune system, invariant networks.

Keywords: anomaly detection, machine learning, fault diagnosis, storage system, root cause analysis.

Citation: Uspenskiy M.B. A survey of the approaches to storage systems fault detection. St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunications and Control Systems, 2019, Vol. 12, No. 4, Pp. 145–158. DOI: 10.18721/JCSTCS.12412

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>)

Введение

Системы хранения данных (СХД) корпоративного уровня в настоящее время являются сложными программно-аппаратными продуктами, которые могут включать в себя, кроме собственно носителей информации, одну или несколько управляющих ЭВМ, дисковые шасси, сетевую инфраструктуру, фабрики (одну или несколько), а также системное программное обеспечение, предоставляющее безопасный доступ к данным, организацию объединения физических носителей информации в логические сущности, управление кластерами, репликацию и избыточность данных. Взаимодействие разнородных программных и аппаратных компонентов в условиях высокой нагрузки может приводить к возникновению сбоев в работе СХД даже в том случае, когда неисправности в отдельных компонентах отсутствуют. Например, в [1, 2] указывается, что на сбои в процессе взаимодействия компонентов приходится до 11 % от общего числа сбоев в работе СХД.

Обнаружение сбоев в процессе функционирования СХД является комплексной задачей, требующей анализа как программных, так и аппаратных компонентов системы, а также процесса их взаимодействия. Актуальность задачи своевременного обнаружения сбоев в работе отдельных программных или аппаратных компонентов СХД и/или СХД в целом определяется тем, что её решение позволяет снизить или устранить вероятность деградации производительности СХД, временной потери доступа к пользовательским данным или потери данных.

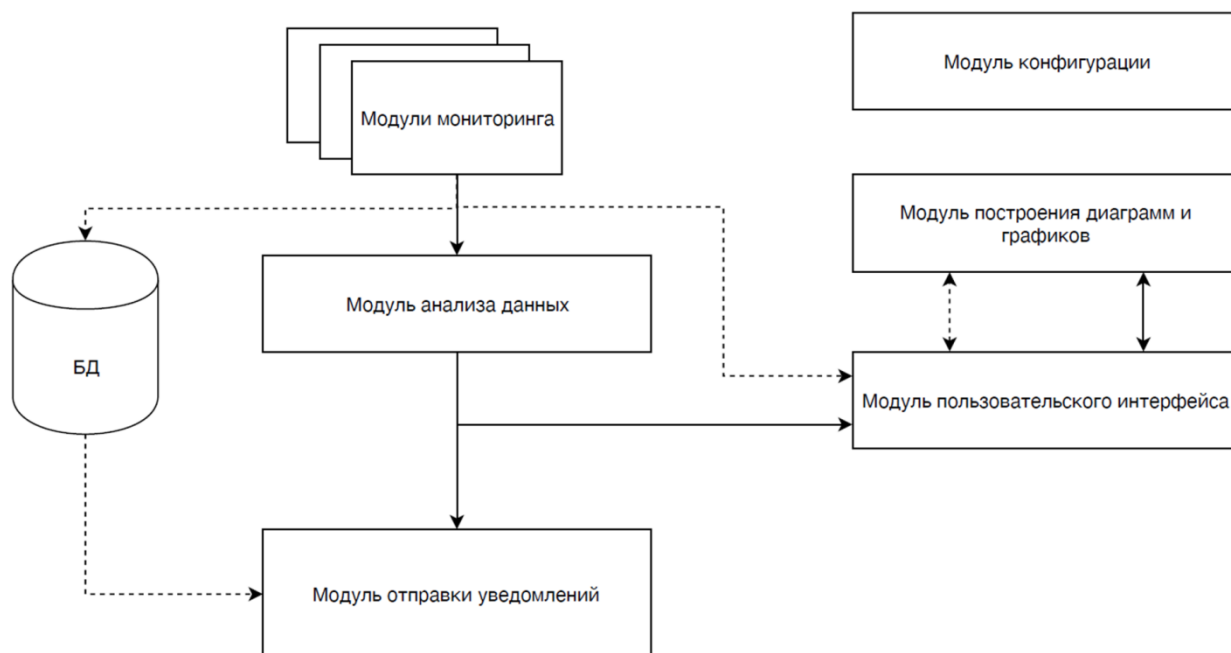
Постановка задачи. Цель исследования — доработка программного комплекса

прогнозирования сбоев систем хранения данных по результатам исследовательских испытаний. В настоящем обзоре предполагается проанализировать существующие подходы к обнаружению сбоев в области компьютерных систем, применяемые в существующих программных решениях для мониторинга и диагностики систем хранения данных, на основании чего выполнить обзор научных публикаций, предлагающих перспективные методы и средства, которые могут быть применены для эффективной реализации данных подходов. Полученная таким образом информация о достоинствах и недостатках этих методов и средств будет применяться в процессе разработки оптимального подхода к обнаружению сбоев для систем хранения данных.

Обзор существующих программных решений в области обнаружения сбоев в работе СХД

Обобщенная структура программного комплекса, полученная на основе анализа представленных на текущий момент решений в области мониторинга и диагностики, приведена на рисунке.

Архитектура, представленная на рисунке, определяется характером решаемых программным комплексом задач: в простейшем случае (связи обозначены пунктирными линиями) программный комплекс не предназначен для интеллектуального анализа данных мониторинга, и обнаружение сбоев осуществляется путем визуального контроля администратором изменений в наблюдаемой СХД. В более сложных случаях используются встроенные средства анализа результатов мониторинга (связи модулей обозначены простыми линиями). Программные комплексы



Обобщенная структура программного комплекса для обнаружения сбоев в работе СХД
The generalized structure of the software package for detecting failures in the storage system

обнаружения сбоев в работе СХД могут включать в себя все или некоторые из программных инструментов, решающих следующие задачи.

- Сбор метрик производительности и/или здоровья — набор модулей мониторинга. Такие модули могут реализовываться, например, в виде программных агентов (см. Zabbix и др.), размещаемых на вычислительных узлах СХД. Отдельный интерес представляет комплекс VirtualWisdom [3], имеющий в своём составе аппаратные средства мониторинга производительности агентов.

- Анализ метрик производительности и/или здоровья для диагностирования или прогнозирования возникновения сбоев — модуль анализа данных. Анализ собранных данных в зависимости от интеллектуальности системы может осуществляться с использованием широкого спектра методов: от сравнения данных мониторинга с пороговыми значениями до использования алгоритмов машинного обучения. Анализ выполняется с целью определения связей (корреляции) между данными мониторинга и состоянием СХД.

- Информирование управляющего персонала о наступлении тех или иных событий, генерируемых модулями мониторинга или анализа данных — модуль отправки уведомлений. События могут быть как заранее заданные разработчиками программного средства, так и определенные администратором. Для решения этой задачи может использоваться отправка уведомлений по электронной почте, смс и/или мобильным мессенджерам.

- Хранение настроек модулей — модуль конфигурации. Такой модуль должен быть связан со всеми прочими модулями системы (эти связи не представлены на рисунке, чтобы не перегружать его лишними деталями). Как правило, могут храниться правила формирования уведомлений, частота сбора данных мониторинга, перечень данных мониторинга, отображаемых диаграмм и графиков и т. д.

- Управление настройками программного комплекса и визуальный контроль за состоянием СХД — модуль графического интерфейса (например, отображение временных графиков изменения зна-

чений параметров с отображением трендов и уровней среднего значения (ElasticStack)).

- Хранение исторических данных о значениях параметров СХД и выходных данных модуля анализа состояний. Хранение осуществляется с помощью базы данных (в некоторых случаях может быть заменена записями в журналы).

Существующие на сегодняшний день программные средства, применяемые для обнаружения сбоев в работе СХД, можно классифицировать по следующим признакам:

- по способу реализации сбора данных: программные (например, Zabbix, ElasticStack) и программно-аппаратные (VirtualWisdom);

- по характеру анализируемых данных: анализ данных мониторинга (Zabbix), анализ журналов (Splunk), анализ данных в общем виде (ElasticStack, Anomaly.io);

- по области применения: средства, применяемые для СХД определенного производителя (встраиваемые в программные средства управления ресурсами конкретных типов СХД — IBM Spectrum Control, Hitachi System Event Management Tool), диагностические средства общего назначения (Zabbix, ElasticStack), средства анализа данных (Anomaly.io). Диагностические средства общего назначения, как правило, предназначены для мониторинга и диагностики серверных ЭВМ в целом, но используются также в задачах мониторинга и обнаружения сбоев в работе СХД.

В [4] описан процесс разработки программного комплекса обнаружения аномалий в реальном времени Anodot, в том числе предложены следующие критерии, которыми необходимо руководствоваться при проектировании программного комплекса для поиска аномалий:

- время реакции на аномалию (может быть требование реального времени или требование к величине временного интервала, в течение которого необходимо обнаружить аномалию);

- объём метрик, которые надо анализировать;

- частота изменения анализируемых временных рядов;

- связность анализируемых метрик, т. е. необходимо ли оценивать аномалии всех метрик в целом или определять их по каждой конкретной метрике;

- устойчивость при отсутствии внешнего контроля (система может опираться на оценки экспертов, а может предъявляться требование).

Для реализации конкретной системы, реализующей устойчивое обнаружение аномалий в большой системе в режиме реального времени с постоянно меняющимися данными и связными метриками, предлагается следующая последовательность шагов при проектировании и обучении: реализация системы универсального сбора метрик, не требующего ручного конфигурирования, исследование нормального поведения объекта анализа (обучение модели при помощи того, что считается нормальным, и вывод статистического теста для отнесения данных к аномальным, если они не описываются моделью), исследование аномального поведения объекта (разработка классификатора аномалий по типу и значимости), исследование связей между метриками с применением алгоритмов кластеризации, точечное обучение на основании оценки распознанных аномалий как ложноположительных или ложноотрицательных.

В рассмотренных программных комплексах выявлены следующие подходы к обнаружению сбоев.

- Обнаружение сбоев на основании заданного списка событий (например, реализовано в программном обеспечении IBM Spectrum Accelerate, Hitachi System Event Management Tool).

- Обнаружение сбоев на основании заданных пороговых значений. Собираемая метрика сравнивается с пороговыми значениями. В случае выхода значения метрики за пределы пороговых значений формируется уведомление администратору (реализовано, например, в Fujitsu ServerView System Monitor).



- Выявление аномалий. В этом случае собираемые метрики анализируются на предмет наличия редких данных, отличающихся от прочих. Решение задачи выявления аномалий может осуществляться разными способами, в том числе с использованием методов машинного обучения (например, методов теории распознавания образов), статистических методов и т. д. Выявление аномалий является основным способом обнаружения сбоев в большинстве существующих программных средств с использованием интеллектуальных средств анализа данных, в том числе Elastic, Anomaly.io и др.

- Сопоставление топологии системы с получаемыми данными мониторинга с определением перечня компонентов в пути прохождения данных.

- Локализация неисправностей с использованием методик анализа коренных причин (root case analysis) на основании данных мониторинга и выявленных аномалий. Реализовано во многих системах, в автоматическом или автоматизированном режимах, например, Infrastructure Analytics Advisor (Hitachi).

Программные продукты, описанные в данном разделе, приведены в сводной таблице. Учитываются следующие характеризующие их признаки:

- тип (анализ данных, система общего назначения, встроенная в продукт конкретного производителя);

- обладает ли система открытым исходным кодом;

- обладает ли система встроенными средствами сбора данных;

Сводная таблица существующих программных средств

Summary table of existing software

Название	Тип	Откры- тый код	Сред- ства сбора данных	Наличие средств локализа- ции неис- правностей	Наличие встроенных средств анализа собранных данных
Anomaly.io	Анализ данных	—	—	—	Визуализация данных, выяв- ление аномалий
Active Health System (HPE)	Встроенные, HPE	—	+	—	—
Active IQ	Встроенные, NetApp	—	+	+	Пороговые значения для формирования уведомлений, визуализация данных
Cacti	Общего назначе- ния, мониторинг	+	+	—	Визуализация данных
CloudIQ	Встроенные, DELL	—	+	+	Выявление аномалий, визуа- лизация данных, построение трендов, оценка рисков
ElasticStack	Анализ данных	+	+	—	Выявление аномалий, поро- говые значения для форми- рования уведомлений, визу- ализация данных
Fujitsu ServerView System Monitor	Встроенные, Fujitsu	—	+	+	—

Окончание таблицы

Название	Тип	Откры- тый код	Сред- ства сбора данных	Наличие средств локализа- ции неис- правностей	Наличие встроенных средств анализа собранных данных
Hitachi System Event Man- agement Tool	Встроенные, Hitachi	—	+	+	Диагностика сбоя по коду события в журнале
Icinga	Общего назна- чения, монито- ринг	—	+	—	Пороговые значения для формирования уведомлений, визуализация данных
IBM Storage Insights	Встроенные, IBM	—	+	+	Пороговые значения для формирования уведомлений, визуализация данных
Nagios	Общего назна- чения, монито- ринг	—	+	—	Пороговые значения для формирования уведомлений, визуализация данных
OnCommand Insight	Встроенные, NetApp	—	+	+	Визуализация данных, поро- говые значения для форми- рования уведомлений, по- строение трендов
Splunk	Анализ данных	—	+	—	Визуализация данных, поро- говые значения для форми- рования уведомлений
Storage Analyt- ics	Встроенные, Dell EMC	—	+	+	Визуализация данных, поро- говые значения для форми- рования уведомлений
VirtualWisdom	Общего назна- чения, монито- ринг	—	+	+	Визуализация данных, поро- говые значения для форми- рования уведомлений
Zabbix	Общего назна- чения, монито- ринг	+	+	—	Визуализация данных, поро- говые значения для форми- рования уведомлений, выяв- ление аномалий

обладает ли система встроенными сред-
ствами локализации неисправностей;

краткое перечисление методов анализа
сбора данных (при возможности их иден-
тификации).

Обзор перспективных подходов к решению задачи выявления аномалий в компьютерных системах

Как было показано в предыдущем раз-
деле, наиболее распространенным методом
обнаружения сбоев является выявление

аномалий. Выявление аномалий обычно
ведется в двух направлениях — обнаруже-
ние выбросов и обнаружение новизны (в
таком случае определяется нормальное по-
ведение системы, и аномалиями считаются
объекты, отсутствующие в обучающей вы-
борке). Существуют различные подходы к
выявлению аномалий: статистические ме-
тоды; методы, основанные на кластериза-
ции и классификации; методы с использо-
ванием искусственных нейронных сетей.
Анализ публикаций, посвященных диагно-



стике сбоев с использованием выявления аномалий за последние пять лет, позволил выделить наиболее перспективные направления в этой области. Далее детально будут рассмотрены следующие группы методов, присутствующие в данных публикациях:

- методы, основанные на кластеризации и классификации;

- методы, основанные на использовании сети инвариантов;

- обнаружение аномалий с применением метода опорных векторов;

- искусственные иммунные системы;

- прочие методы, в том числе, использующие подходы статистического анализа.

Методы, основанные на кластеризации и классификации. Методы обнаружения аномалий, основанные на решении задач классификации или кластеризации, являются наиболее распространенными в современных работах, посвященных обнаружению сбоев в области диагностики вычислительных устройств или устройств хранения данных.

В [5] предложен подход к диагностике сбоев в реальном времени, основанный на рекурсивной оценке плотности. Оценка плотности опирается на распределение Коши, параметры которого можно обновлять рекурсивно, что позволяет хранить в памяти минимальное количество данных. Преимуществом предложенного подхода является отсутствие требования по наличию модели диагностируемого процесса или исторических данных для формирования обучающей выборки.

В работе [6] изучено использование адаптивной ядерной оценки плотности для определения выбросов в нелинейных системах. В соответствии с этим подходом каждой выборке данных назначается локальное значение выброса, показывающее, насколько одна выборка отличается от других в ее окрестности.

Процесс организации анализа сетевого потока данных на основе кластеризованных образов рассмотрен в [7]. Метод направлен на решение задачи обработки большого объема исходных данных (десять-

ки миллионов пакетов в секунду на одно сетевое соединение в 10 Гб, что не позволяет прямо применять детальный анализ) путем агрегирования переменных, описывающих трафик в абстрактные образы (для кластеризации в работе используется метод k -средних), после чего задача определения аномалий сводится к задаче сравнения образов.

В работе [8] предложен комбинированный метод диагностики, использующий совместно модельный подход к диагностике и классификаторы аномалий для локализации неисправностей, направленный на совмещение преимуществ обоих подходов и повышение точности реализованного алгоритма путем переобучения классификаторов на основании выявленных в процессе эксплуатации сбоев. Диагностирование происходит по следующей схеме: на основании модельного подхода определяются потенциальные сбои, оценка вероятности которых осуществляется при помощи классификаторов аномалий. С помощью данного подхода решены задачи определения и локализации неисправностей в условиях недостатка данных для обучающих выборок, переобучения классификаторов в автоматическом режиме на основании выявленных сбоев и поиска множественных сбоев в условиях наличия только единичных сбоев в обучающих выборках.

Прикладной случай решения задачи обнаружения сбоев на основании метрик производительности компонентов виртуальных машин с использованием платформы анализа больших наборов данных Hadoop MapReduce совместно с наивным классификатором Байеса в сервисе предоставления виртуальных машин и облачных хранилищ рассмотрен в [9]. Точность полученного метода достигает 89,8 % с ростом объема поступающих данных.

Масштабируемый непараметрический метод, предназначенный для поиска аномалий производительности в вычислительных системах большого размера, предложен в [10]. Для решения данной задачи представлен децентрализованный подход,

основанный на широко распространенном методе сравнения узлов (описан, например, в [11]), включающий в себя следующие этапы: иерархическая группировка системных узлов, непараметрическая кластеризация и двухступенчатое определение аномальных узлов в каждой группе. Преимущества такого подхода — хорошая масштабируемость и высокая производительность полученного решения.

Сети инвариантов. В последнее время широкое распространение получили методы, основанные на комплексном описании поведения систем с использованием сетей инвариантов. В сетях инвариантов узел представляет собой компонент системы, а дуга — значимое, стабильное взаимодействие между двумя компонентами. Инвариантная модель ориентирована на определение стабильных значимых зависимостей между парами компонентов, которые наблюдаются при помощи записи временных рядов, для определения состояния системы. Сильная связь между компонентами называется *инвариантной (корреляционной) зависимостью*. Объединяя инварианты, определенные по всем наблюдаемым компонентам, можно получить глобальный профиль системы. В таком случае аномальное поведение системы и источник аномалии определяется на основании поиска поврежденных инвариантов (исчезающие корреляции).

Подход к определению неисправностей в программном обеспечении, использующий рассогласование динамических инвариантов, предложен в [12]. Подход направлен на решение таких проблем динамических инвариантов, как высокая вычислительная сложность определения инвариантов, фильтрации ложноположительных срабатываний и устранение избыточности. Смысл подхода заключается в предварительной фильтрации подозрительных функций и последовательном применении к ним инструментов определения инвариантов.

Рассмотренный в [13] подход по определению причинно-следственных анома-

лий при помощи сетей инвариантов с использованием временного и динамического анализа исчезающих корреляций направлен на решение таких проблем метода сетей инвариантов, как невозможность определения маршрута распространения сбоя по сети, наличия узлов с максимальным процентом исчезающих корреляций, не являющихся при этом корневыми случайными аномалиями, отсутствие возможности использовать априорные знания об аномальных узлах и временных параметрах исчезающих корреляций. В качестве инструментов для решения этих проблем предлагается сетевая диффузия для моделирования распространения причинно-следственных аномалий, применение оценок маршрута распространяющихся причинно-следственных аномалий для реконструкции исчезающих корреляций, а также использование стратегии нормализации, позволяющей не учитывать экстремальные значения или выбросы без необходимости явно исключать их из набора данных. Кроме того, алгоритм позволяет использовать априорное знание для оценки состояния отдельных аномальных узлов в некоторые моменты времени, с поправкой на то, что такое априорное знание может быть сильно зашумленным.

Метод опорных векторов. Давно используемый для задач классификации метод опорных векторов также часто применяется для выявления аномалий.

В исследовании [14] описан случай применения метода опорных векторов для определения аномальных виртуальных машин в облачной среде с целью предотвращения деградации производительности. Для этих целей авторами разработана программная платформа EaAD, позволяющая определять аномалии в виртуальных машинах с учетом влияния внешней среды. Для оценки виртуальных машин используются их метрики производительности. Для определения аномалий в EaAD реализуются алгоритмы, основанные на различных вариантах метода опорных векторов.



Модификация метода опорных векторов для одного класса, направленная на решение проблемы уязвимости метода к выбросам в обучающей выборке, предложена в [15]. Разработанный устойчивый метод опорных векторов для одного класса подразумевает подавление возникающих выбросов в обучающей выборке за счет введения отрицательных весовых коэффициентов.

Искусственные иммунные системы. Искусственные иммунные системы — это вычислительные системы, основанные на принципах работы иммунной системы. Они включают в себя в том числе негативный алгоритм отбора, иммунный сетевой алгоритм и пр. Использование иммунных систем в качестве основы для алгоритмов выявления аномалий обосновывается тем, что биологические иммунные системы, так же как и алгоритмы выявления аномалий, нацелены на выявление информации, отличной от нормальной [16, 17].

В работе [18] предложено использовать для обнаружения сбоев две разновидности алгоритма негативного отбора — FB-NSA (Fixed Boundary NSA) и его доработанный вариант FFB-NSA (Fine Fixed Boundary NSA). Принципиальное отличие от прочих работ в области алгоритмов отрицательного отбора (например, в [19] предложен детектор в виде гиперкуба, в [20] — детектор в виде гиперэллипсоида, в [21] — детектор с множественной формой) заключается в направленности не на исследование формы детекторов, а на то, чтобы добиться их неизменности. Константный детектор зависит только от обучающей выборки и не связан с временем обучения. Оба предложенных алгоритма генерируют слой детекторов, расположенных между нормальными и аномальными обучающими данными. При этом детекторы имеют постоянное число, размер и положение. Доработанный алгоритм включает в себе механизм охвата слепой зоны, возникающей при использовании больших детекторов и с малым значением параметра m (размера тестовой выборки).

Прочие методы. В работе [22] предложено выполнять обнаружение сбоев и определение источника их возникновения на основе графической вероятностной модели. Метод предлагает альтернативу использования традиционной Байесовой сети в качестве графической модели в виде модели причинно-следственных связей системы, с применением ядерной оценки плотности для оценки вероятностной функции плотности вместо изучения параметров, характерного для Байесовой сети.

Подход к диагностике жестких дисков на основании полупараметрических моделей и статистических оценок рассмотрен в [23]. Для моделирования распределения параметров SMART исправного жесткого диска используется смесь Гауссовских распределений. Аномалиями при этом считаются случаи, когда рассчитанная статистическая оценка различий между поведением диска и полученной модели превышает пороговое значение. В работе [24] рассмотрено сравнение алгоритма, основанного на смеси Гауссовских распределений с несколькими алгоритмами, в том числе, алгоритмом, основанным на использовании расстояния Махалнобиса, алгоритмом, основанным на методе опорных векторов, и стандартным алгоритмом диагностики SMART, и демонстрируется его преимущество в точности. Для оценки качества работы алгоритмов используются метрики FDR (fault detection rate) и FAR (false alarm rate) на реальном наборе статистических данных о сбоях жестких дисков. Алгоритм, основанный на смеси Гауссовских распределений, оказывается наиболее эффективным, с FDR 80,59 % и с FAR 0 %;

Алгоритм, основанный на латентной корреляционной вероятностной модели, предназначенный для поиска аномалий в данных мониторинга состояния оборудования, представлен в [25]. Предложенный алгоритм направлен на решение основных проблем, возникающих в процессе поиска аномалий в данных мониторинга: необходимости обработки большого объема дан-

ных, генерируемых системами мониторинга, сложности идентификации источника аномалии (т. к. в общем случае аномалия в одном компоненте оборудования влияет на данные мониторинга сразу нескольких связанных с ним компонентов) и наличия шума в данных мониторинга. Для решения этих проблем вводится понятие латентной корреляции, которая определяет связь между различными наборами данных мониторинга в некоторый момент времени.

Подход к обнаружению сбоев для одного класса, основанный на использовании генеративно-состязательных сетей, предложен в [26]. Генеративно-состязательная сеть является алгоритмом машинного обучения без учителя, в котором одна нейронная сеть генерирует обучающие выборки для другой нейронной сети, пытающейся выявить в созданных выборках аномалии (см., например, описание генеративной сети в [27]). В подходе, описанном в [26], первая нейронная сеть пытается определить нормальное протекание процесса, а вторая принимает окончательное решение о наличии аномалий. Для обеспечения конвергентности системы в процессе обучения и для повышения точности классификатора предложено использовать специализированный алгоритм, проверяющий обученные модели на специальном валидационном наборе данных. Полученный подход позволяет реализовать алгоритм более производительный, чем классические методы опорных векторов и изолирующего леса.

Прикладная задача выявления ранее не выявленных аномалий в полётных данных самолетов, собранных в единой базе данных, рассмотрена в [28]. Для решения данной задачи предложено использовать комбинированный подход с применением методов формализации экспертного знания и методов машинного обучения без учителя. Полученный в итоге алгоритм имеет неспециализированный характер и применим к выявлению аномалий в различных промышленных системах. Для решения проблемы большого объёма анализируемых данных используется вейвлет-

преобразование, позволяющее преобразовать непрерывный временной ряд в набор дискретных отсчетов, сохраняющих пространственно-временные характеристики сигнала, после чего этот набор сворачивается в матрицу сходства, к которой затем применяется алгоритм иерархической кластеризации, позволяющий построить набор кластеров, каждый из которых содержит схожие по параметрам сведения о полетах. Далее определяются параметры, позволяющие получить максимальное расхождение между кластерами, после чего данные анализируются экспертами и применяются для обучения алгоритмов с учителем.

Обнаружение аномалий в условиях смешанных численно-категориальных исходных данных большого масштаба изучено в [29]. Для решения этой задачи предложена комбинированная вероятностная мера, представляющая собой маргинальную плотность для непрерывных переменных и условную вероятность для категориальных переменных. В результате итоговый алгоритм, базирующийся на модели, обученной с помощью целевой функции максимального сходства и оптимизированной с помощью стохастического градиентного спуска, демонстрирует наиболее высокую точность среди всех алгоритмов определения аномалий по нескольким наборам параметров, а также максимальную производительность и масштабируемость.

Заключение

В ходе анализа актуальных тенденций в области диагностики систем хранения данных корпоративного уровня рассмотрены наиболее распространенные программные системы управления и мониторинга СХД, анализа данных и программные средства общего назначения. На текущий момент можно сделать вывод о том, что почти все рассмотренные актуальные программные средства предлагают некоторую функциональность для автоматического, автоматизированного или выполняемого администратором поиска неисправностей варьируемой сложности [30, 31].



Кроме того, можно сделать вывод о том, что системы общего назначения, предоставляющие унифицированную функциональность по сбору, обработке и визуализации информации, получают всё большее распространение. Преимуществом таких систем является использование ими большого количества разнообразных алгоритмов выявления аномалий, в том числе вынесенных в сферу облачного вычисления, что позволяет экономить вычислительные ресурсы СХД. При этом алгоритмы выявления аномалий сами по себе не обладают знаниями о топологии и устройстве конкретных систем хранения данных и, следовательно, в общем случае не подходят для определения типов возникающих неисправностей, пути распространения ошибки по топологии системы и конкретных компонентов, в которых возникла неисправность. Программные средства, предоставляемые производителями систем хранения данных, имеют более слабое алгоритмическое обеспечение по выявлению аномалий, но за счет наличия у разработчиков экспертного знания о предмете анализа, обладают, как правило, инструментами локализации неисправности.

Анализ публикаций, посвященных обнаружению неисправностей или выявлению аномалий, позволяет сделать вывод о том, что в настоящее время широкое распространение получают алгоритмы, комбинирующие различные группы методов

выявления аномалий. Задача повышения точности алгоритмов во многих случаях отходит на второй план, делая более приоритетными задачи повышения производительности, экономии системных ресурсов, масштабируемости и т. д. В рассмотренных публикациях применяются как классические группы методов, такие как методы на основе опорных векторов, методы, основанные на алгоритмах кластеризации и классификации и т. д., так и более редкие, представляющие, возможно, больший интерес, такие как алгоритмы искусственных иммунных сетей.

Для разработки подходов к обнаружению и локализации неисправностей в работе СХД перспективной представляется такая организация системы обнаружения сбоев, при которой возможно комбинировать формализованное экспертное знание о предметной области с получаемой извне информацией о наличии аномалий в данных мониторинга, т. е., фактически, параллельное применение методов, основанных на модельно-ориентированном подходе и подходе, ориентированном на анализе массива данных.

Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы». Уникальный идентификатор RFMEFI58117X0023.

СПИСОК ЛИТЕРАТУРЫ

1. Hewlett Packard Enterprise Development LP. Can Machine Learning Prevent Application Downtime? // URL: <https://cdm-cdn.nimblestorage.com/2017/08/23090218/Can-Machine-Learning-White-Paper-1708-FINAL-print.pdf> (Дата обращения: 04.04.2019).
2. Wang D. Artificial intelligence makes flash storage predictive // URL: <https://www.hpe.com/us/en/insights/articles/artificial-intelligence-makes-flash-storage-predictive-1803.html> (Дата обращения: 01.04.2019).
3. Lelii S. VirtualWisdom adds storage probe for NetApp NAS array // URL: <https://searchstorage.techtarget.com/news/2240219781/VirtualWisdom-adds-storage-probe-for-NetApp-NAS-array> (Дата обращения: 20.04.2019).
4. Toledano M., Cohen I., Ben-Simhon Y., Tadeski I. Real-time anomaly detection system for time series at scale // Proc. of the KDD 2017: Workshop on Anomaly Detection in Finance. 2018. Pp. 56–65.
5. Costa B.S.J., Angelov P.P., Guedes L.A. Real-time fault detection using recursive density estimation // J. of Control, Automation and Electrical Systems. 2014. Vol. 4. No. 25. Pp. 428–437. DOI: 10.1007/s40313-014-0128-4
6. Zhang L., Lin J., Karim R. Adaptive kernel density-based anomaly detection for nonlinear systems // Knowledge-Based Systems. 2018. No. 139. Pp. 50–63. DOI: 10.1016/j.knosys.2017.10.009
7. Kim J., Sim A., Tierney B., Suh S., Kim I. Multivariate network traffic analysis using clustered

- patterns // Computing. 2018. Vol. 101. No. 4. Pp. 339–361. DOI:10.1007/s00607-018-0619-4
8. **Jung D., Ng K.Y., Frisk E., Krysanter M.** Combining model-based diagnosis and data-driven anomaly classifiers for fault isolation // Control Engineering Practice. 2018. No. 80. Pp. 146–156. DOI: 10.1016/j.conengprac.2018.08.013
9. **Alkasem A., Liu H., Shafiq M.** Improving fault diagnosis performance using Hadoop MapReduce for efficient classification and analysis of large data sets // J. of Computers. 2018. Vol. 29. No. 4. Pp. 185–202.
10. **Yu L., Lan Z.** A scalable, non-parametric method for detecting performance anomaly in large scale computing // IEEE Transactions on Parallel and Distributed Systems. 2016. Vol. 27. No. 7. Pp. 1902–1914.
11. **Lan Z., Zheng Z., Li Y.** Toward automated anomaly identification in large-scale systems // IEEE Trans. Parallel Distrib. Syst. 2010. Vol. 21. Pp. 174–187.
12. **Wang X., Liu Y.** Fault localization using disparities of dynamic invariants // J. of Systems and Software. 2016. Vol. 122. Pp. 144–154. DOI: 10.1016/j.jss.2016.09.014
13. **Cheng W., Ni J., Zhang K., Chen H., Jiang G., Shi Y., Wang W.** Ranking causal anomalies for system fault diagnosis via temporal and dynamical analysis on vanishing correlations // ACM Transactions on Knowledge Discovery from Data. 2017. Vol. 11. No. 4. Pp. 1–28. DOI:10.1145/3046946
14. **Wang G.P., Wang J.W.** An anomaly detection framework for detecting anomalous virtual machines under cloud computing environment // Internat. J. of Security and Its Applications. 2016. Vol. 10. No. 1. Pp. 75–86.
15. **Yin S., Zhu X., Jing C.** Fault detection based on a robust one class support vector machine // Neurocomputing. 2014. Vol. 145. Pp. 263–268. DOI: 10.1016/j.neucom.2014.05.035
16. **Dasgupta D., González F.** An immunity-based technique to characterize intrusions in computer networks // IEEE Trans. Evol. Comput. 2002. Vol. 6. Pp. 281–291.
17. **Silva G.C., Caminhas W.M., Palhares R.M.** Artificial immune systems applied to fault detection and isolation: A brief review of immune response-based approaches and a case study // Applied Soft Computing. 2017. Vol. 57. Pp. 118–131.
18. **Li D., Liu S., Zhang H.** Negative selection algorithm with constant detectors for anomaly detection // Applied Soft Computing. 2015. Vol. 36. Pp. 618–632.
19. **González F., Gyme J., Dasgupta D.** An evolutionary approach to generate fuzzy anomaly (attack) signatures // Proc. of IEEE Systems, Man and Cybernetics Society. IEEE, 2003. Pp. 251–259.
20. **Shapiro J.M., Lamont G.B., Peterson G.L.** An evolutionary algorithm to generate ellipsoid network intrusion detectors // Proc. of the 2005 Workshops on Genetic and Evolutionary Computation. ACM, 2005. Pp. 178–180.
21. **Balachandran S., Dasgupta D., Nino F., Garrett D.** A framework for evolving multi-shaped detectors in negative selection // Proc. of IEEE Symp. on Foundations of Computational Intelligence. IEEE, 2007. Pp. 401–408.
22. **Chen X., Wang J., Zhou J.** Fault detection and backtrace based on graphical probability Model // Prognostics and System Health Management Conf. (PHM-Chongqing). IEEE, 2018. Pp. 584–590.
23. **Queiroz L.P., et al.** Fault detection in hard disk drives based on a semi parametric model and statistical estimators // New Generation Computing. 2018. Vol. 36. No. 1. Pp. 5–19.
24. **Queiroz L.P., et al.** Fault detection in hard disk drives based on mixture of Gaussians // 5th Brazilian Conf. on Intelligent Systems. IEEE, 2016. Pp. 145–150.
25. **Ding J., et al.** An anomaly detection approach for multiple monitoring data series based on latent correlation probabilistic model // Applied Intelligence. 2016. Vol. 44. No. 2. Pp. 340–361.
26. **Plakias S., Boutalis Y.S.** Exploiting the generative adversarial framework for one-class multi-dimensional fault detection // Neurocomputing. 2019. Vol. 332. Pp. 396–405.
27. **Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y.** Generative Adversarial Networks. 2016. arXiv:1406.2661[stat.ML]
28. **Mack D.L.C., Biswas G., Khorasgani H., Mylaraswamy D., Bharadwaj R.** Combining expert knowledge and unsupervised learning techniques for anomaly detection in aircraft flight data // Automatisierungstechnik. 2018. Vol. 66. No. 4. Pp. 291–307. DOI:10.1515/auto-2017-0120
29. **Eiras-Franco C., et al.** Large scale anomaly detection in mixed numerical and categorical input spaces // Information Sciences. 2019. Vol. 487. Pp. 115–127.
30. **Higley L.** Storage analytics: Can we put any more lipstick on that pig? // URL: <https://cloud.kapostcontent.net/pub/3da21605-fc17-4712-991a-1c49dc77b871/mfx131e-pc-mon-130-higleyl.pdf?kui=xxPHjAO870Nzv0HTEjftEw> (Дата обращения: 10.04.2019).
31. **Gopisetty S.** Evolution of storage management: Transforming raw data into information // IBM Journal of Research and Development. 2008. Vol. 52. No 4.5. Pp. 341–352.

Статья поступила в редакцию 10.05.2019.



REFERENCES

1. Hewlett Packard Enterprise Development LP. Can Machine Learning Prevent Application Downtime? Available: <https://cdm-cdn.nimblestorage.com/2017/08/23090218/Can-Machine-Learning-White-Paper-1708-FINAL-print.pdf> (Accessed: 04.04.2019).
2. Wang D. Artificial intelligence makes flash storage predictive. Available: <https://www.hpe.com/us/en/insights/articles/artificial-intelligence-makes-flash-storage-predictive-1803.html> (Accessed: 01.04.2019).
3. Lelii S. VirtualWisdom adds storage probe for NetApp NAS array. Available: <https://searchstorage.techtarget.com/news/2240219781/VirtualWisdom-adds-storage-probe-for-NetApp-NAS-array> (Accessed: 20.04.2019).
4. Toledano M., Cohen I., Ben-Simhon Y., Tadeski I. Real-time anomaly detection system for time series at scale. *Proceedings of the KDD 2017: Workshop on Anomaly Detection in Finance*, 2018, Pp. 56–65.
5. Costa B.S.J., Angelov P.P., Guedes L.A. Real-Time Fault Detection Using Recursive Density Estimation. *Journal of Control, Automation and Electrical Systems*, 2014, Vol. 4, No. 25, Pp. 428–437. DOI:10.1007/s40313-014-0128-4
6. Zhang L., Lin J., Karim R. Adaptive kernel density-based anomaly detection for nonlinear systems. *Knowledge-Based Systems*, 2018, No. 139, Pp. 50–63. DOI: 10.1016/j.knosys.2017.10.009
7. Kim J., Sim A., Tierney B., Suh S., Kim I. Multivariate network traffic analysis using clustered patterns. *Computing*, 2018, Vol. 101, No. 4, Pp. 339–361. DOI:10.1007/s00607-018-0619-4
8. Jung D., Ng K.Y., Frisk E., Krysander M. Combining model-based diagnosis and data-driven anomaly classifiers for fault isolation. *Control Engineering Practice*, 2018, No. 80, Pp. 146–156. DOI: 10.1016/j.conengprac.2018.08.013
9. Alkasem A., Liu H., Shafiq M. Improving fault diagnosis performance using Hadoop MapReduce for efficient classification and analysis of large data sets. *Journal of Computers*, 2018, Vol. 29, No. 4, Pp. 185–202.
10. Yu L., Lan Z. A scalable, non-parametric method for detecting performance anomaly in large scale computing. *IEEE Transactions on Parallel and Distributed Systems*, 2016, Vol. 27, No. 7, Pp. 1902–1914.
11. Lan Z., Zheng Z., Li Y. Toward automated anomaly identification in large-scale systems. *IEEE Trans. Parallel Distrib. Syst.*, 2010, Vol. 21, Pp. 174–187.
12. Wang X., Liu Y. Fault localization using disparities of dynamic invariants. *Journal of Systems and Software*, 2016, Vol. 122, Pp. 144–154. DOI: 10.1016/j.jss.2016.09.014
13. Cheng W., Ni J., Zhang K., Chen H., Jiang G., Shi Y., Wang W. Ranking causal anomalies for system fault diagnosis via temporal and dynamical analysis on vanishing correlations. *ACM Transactions on Knowledge Discovery from Data*, 2017, Vol. 11, No. 4, Pp. 1–28. DOI:10.1145/3046946
14. Wang G.P., Wang J.W. An anomaly detection framework for detecting anomalous virtual machines under cloud computing environment. *International Journal of Security and its Applications*, 2016, Vol. 10, No. 1, Pp. 75–86.
15. Yin S., Zhu X., Jing C. Fault detection based on a robust one class support vector machine. *Neurocomputing*, 2014, Vol. 145, Pp. 263–268. DOI: 10.1016/j.neucom.2014.05.035
16. Dasgupta D., González F. An immunity-based technique to characterize intrusions in computer networks. *IEEE Trans. Evol. Comput.*, 2002, Vol. 6, Pp. 281–291.
17. Silva G.C., Caminhas W.M., Palhares R.M. Artificial immune systems applied to fault detection and isolation: A brief review of immune response-based approaches and a case study. *Applied Soft Computing*, 2017, Vol. 57, Pp. 118–131.
18. Li D., Liu S., Zhang H. Negative selection algorithm with constant detectors for anomaly detection. *Applied Soft Computing*, 2015, Vol. 36, Pp. 618–632.
19. González F., Gymez J., Dasgupta D. An evolutionary approach to generate fuzzy anomaly (attack) signatures. *Proceedings of IEEE Systems, Man and Cybernetics Society*, IEEE, 2003, Pp. 251–259.
20. Shapiro J.M., Lamont G.B., Peterson G.L. An evolutionary algorithm to generate ellipsoid network intrusion detectors. *Proceedings of the 2005 Workshops on Genetic and Evolutionary Computation*, ACM, 2005, Pp. 178–180.
21. Balachandran S., Dasgupta D., Nino F., Garrett D. A framework for evolving multi-shaped detectors in negative selection. *Proceedings of IEEE Symposium on Foundations of Computational Intelligence*, IEEE, 2007, Pp. 401–408.
22. Chen X., Wang J., Zhou J. Fault detection and backtrace based on graphical probability model. *Prognostics and System Health Management Conference (PHM-Chongqing)*, IEEE, 2018, Pp. 584–590.
23. Queiroz L.P., et al. Fault detection in hard disk drives based on a semi parametric model and statistical estimators. *New Generation Computing*, 2018, Vol. 36, No. 1, Pp. 5–19.

24. **Queiroz L.P., et al.** Fault detection in hard disk drives based on mixture of Gaussians. *5th Brazilian Conference on Intelligent Systems*, IEEE, 2016, Pp. 145–150.
25. **Ding J., et al.** An anomaly detection approach for multiple monitoring data series based on latent correlation probabilistic model, *Applied Intelligence*, 2016, Vol. 44, No. 2, Pp. 340–361.
26. **Plakias S., Boutalis Y.S.** Exploiting the generative adversarial framework for one-class multi-dimensional fault detection. *Neurocomputing*, 2019, Vol. 332, Pp. 396–405.
27. **Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y.** *Generative Adversarial Networks*. 2016. arXiv:1406.2661[stat.ML]
28. **Mack D.L.C., Biswas G., Khorasgani H., Mylaraswamy D., Bharadwaj R.** Combining expert knowledge and unsupervised learning techniques for anomaly detection in aircraft flight data. *Automatisierungstechnik*, 2018, Vol. 66, No. 4, Pp. 291–307. DOI:10.1515/auto-2017-0120
29. **Eiras-Franco C., et al.** Large scale anomaly detection in mixed numerical and categorical input spaces. *Information Sciences*, 2019, Vol. 487, Pp. 115–127.
30. **Higley L.** Storage analytics: Can we put any more lipstick on that pig? Available: <https://cloud.kapostcontent.net/pub/3da21605-fc17-4712-991a-1c49dc77b871/mfx131e-pc-mon-130-higleyl.pdf?kui=xxPHjAO870Nzv0HTEjftEw> (Accessed: 10.04.2019).
31. **Gopisetty S.** Evolution of storage management: Transforming raw data into information. *IBM Journal of Research and Development*, 2008, Vol. 52, No 4.5, Pp. 341–352.

Received 10.05.2019.

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

УСПЕНСКИЙ Михаил Борисович

USPENSKIY Mikhail B.

E-mail: umihail@list.ru

НАУЧНОЕ ИЗДАНИЕ
«НАУЧНО-ТЕХНИЧЕСКИЕ ВЕДОМОСТИ
САНКТ-ПЕТЕРБУРГСКОГО
ГОСУДАРСТВЕННОГО ПОЛИТЕХНИЧЕСКОГО УНИВЕРСИТЕТА.
ИНФОРМАТИКА. ТЕЛЕКОММУНИКАЦИИ. УПРАВЛЕНИЕ»

«ST. PETERSBURG STATE POLYTECHNICAL UNIVERSITY JOURNAL.
COMPUTER SCIENCE. TELECOMMUNICATIONS AND CONTROL SYSTEMS»

Том 12, № 4, 2019

Учредитель — Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Санкт-Петербургский государственный политехнический университет»

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор).
Свидетельство о регистрации ПИ № ФС77-51457 от 19.10.2012 г.

Редакция журнала

д-р техн. наук, профессор *А.С. Коротков* — главный редактор
Е.А. Калинина — литературный редактор, корректор
Г.А. Пышкина — ответственный секретарь, выпускающий редактор

Телефон редакции (812)552-62-16

E-mail: infocom@spbstu.ru

Компьютерная верстка *Е.А. Корнукова*

Лицензия ЛР № 020593 от 07.08.97

Дата выхода 27.12.2019. Формат 60×84 1/8

Санкт-Петербургский политехнический университет Петра Великого
Адрес университета и редакции: 195251, Санкт-Петербург, ул. Политехническая, д. 29.

УСЛОВИЯ ПУБЛИКАЦИИ СТАТЕЙ

в журнале «Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление»

1. ОБЩИЕ ПОЛОЖЕНИЯ

Журнал «Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Телекоммуникации. Управление» является периодическим научным рецензируемым изданием. Зарегистрировано Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации ПИ № ФС77-51457 от 19 октября 2012 г. С 2008 года выпускается в составе сериального периодического издания «Научно-технические ведомости СПбГПУ» (ISSN 1994-2354).

Издание с 2002 года входит в Перечень ведущих научных рецензируемых журналов и изданий (перечень ВАК) и принимает материалы научных исследований, а также статьи для опубликования основных результатов диссертаций на соискание ученой степени доктора наук и кандидата наук по следующим основным научным направлениям: **ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА, РАДИОТЕХНИКА И СВЯЗЬ, ЭЛЕКТРОНИКА, ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА, УПРАВЛЕНИЕ В СОЦИАЛЬНЫХ И ЭКОНОМИЧЕСКИХ СИСТЕМАХ**. Научные направления журнала учитываются ВАК Минобрнауки РФ при защите докторских и кандидатских диссертаций в соответствии с Номенклатурой специальностей научных работников.

Сведения о публикациях представлены в РИНЦ, в Реферативном журнале ВИНИТИ РАН, в международной справочной системе «Ulrich's Periodical Directory».

Периодичность выхода журнала — 4 номера в год.

2. ТРЕБОВАНИЯ К ПРЕДОСТАВЛЯЕМЫМ МАТЕРИАЛАМ

2.1. Оформление материалов

1. Рекомендуемый объем статей для авторов с ученой степенью доктора наук, званием профессора, соискателей ученой степени доктора наук (докторантов) 12–20 страниц формата А-4 с учетом графических вложений. Количество графических вложений (диаграмм, графиков, рисунков, таблиц, фотографий и т. п.) не должно превышать 4.

2. Рекомендуемый объем статей для преподавателей, авторов без ученой степени, соискателей ученой степени кандидата наук — 8–15 страниц формата А-4; аспирантов — 8 страниц формата А-4 с учетом графических вложений. Количество графических вложений (диаграмм, графиков, рисунков, таблиц, фотографий и т. п.) не должно превышать 3.

3. Авторы должны придерживаться следующей обобщенной структуры статьи: вводная часть (0,5–1 стр., актуальность, существующие проблемы); основная часть (постановка и описание задачи, изложение и суть основных результатов); заключительная часть (0,5–1 стр., предложения, выводы), список литературы (оформление по ГОСТ 7.05.-2008).

4. Число авторов статьи не должно превышать трех человек.

5. Набор текста осуществляется в редакторе **MS Word**, формул — в редакторе **MathType**. Таблицы набираются в том же формате, что и основной текст.

6. Шрифт — TNR, размер шрифта основного текста — 14, интервал — 1,5; таблицы большого размера могут быть набраны 12 кеглем. Параметры страницы: поля слева — 3 см, сверху, снизу — 2,5 см, справа — 2 см, текст размещается без переносов. Абзацный отступ — 1 см.

2.2. Предоставление материалов

Вместе с материалами статьи должны быть обязательно предоставлены:

- номер УДК в соответствии с классификатором (в заголовке статьи);
- аннотация на русском и английском языках;
- ключевые слова (5–7) на русском и английском языках;
- сведения об авторах на русском и английском языках: ФИО, место работы, должность, ученое звание, ученая степень, контактные телефоны, e-mail;
- аспиранты представляют документ отдела аспирантуры, заверенный печатью;
- акт экспертизы о возможности опубликования материалов в открытой печати.

С авторами статей заключается издательский лицензионный договор.

Предоставление всех материалов осуществляется в электронном виде через личный кабинет **ЭЛЕКТРОННОЙ РЕДАКЦИИ** по адресу <http://journals.spbstu.ru>

2.3. Рассмотрение материалов

Предоставленные материалы (п. 2.2) первоначально рассматриваются редакционной коллегией и передаются для рецензирования. После одобрения материалов, согласования различных вопросов с автором (при необходимости) редакционная коллегия сообщает автору решение об опубликовании статьи. В случае отказа в публикации статьи редакция направляет автору мотивированный отказ.

При отклонении материалов из-за нарушения сроков подачи, требований по оформлению или как не отвечающих тематике журнала материалы не публикуются и не возвращаются.

Редакционная коллегия не вступает в дискуссию с авторами отклоненных материалов.

Публикация материалов аспирантов очной бюджетной формы обучения осуществляется бесплатно в соответствии с очередностью.

При поступлении в редакцию значительного количества статей их прием в очередной номер может закончиться **ДОСРОЧНО**.

Более подробную информацию можно получить:

на сайте журнала <http://ntv.spbstu.ru>

по телефону редакции +7(812) 552-62-16 с 10⁰⁰ до 18⁰⁰ Галина Александровна

или по e-mail: infocom@spbstu.ru