

**СБОРНИК ТРУДОВ
ХМЕЖДУНАРОДНОЙ
НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ
“ПРОГРАММНАЯ ИНЖЕНЕРИЯ И КОМПЬЮТЕРНАЯ ТЕХНИКА
(МАЙОРОВСКИЕ ЧТЕНИЯ)”
“PROCEEDINGS OF THE MAJOROV INTERNATIONAL CONFERENCE
ON SOFTWARE ENGINEERING AND COMPUTER SYSTEMS ”**



Санкт-Петербург
2019

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

Сборник трудов
X международной научно-практической
конференции
«Программная инженерия и компьютерная техника
(Майоровские чтения)»



Санкт-Петербург
2019

Сборник трудов X международной научно-практической конференции «Программная инженерия и компьютерная техника (Майоровские чтения)» /
Под ред. д.т.н., проф. Т.И. Алиева. СПб: Университет ИТМО, 2019. – 113 с.

Сборник содержит статьи студентов, аспирантов и сотрудников, отобранные по результатам докладов, представленных на X международной научно-практической конференции «Программная инженерия (Майоровские чтения)», состоявшейся в 2018 г. Целью конференции является ознакомление научной общественности с результатами исследований, выполненных студентами, аспирантами и сотрудниками российских и иностранных университетов в рамках научного направления «Программная инженерия и компьютерная техника», а также в ходе выполнения научно-исследовательских и опытно-конструкторских работ, проводимых по заказу предприятий и организаций Российской Федерации.



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО – становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО,
Авторы,
2019

ISBN: 978-5-7577-0602-3

Сборник трудов
X международной научно-практической
конференции молодых ученых
«Программная инженерия и компьютерная
техника (Майоровские чтения)»

СОДЕРЖАНИЕ

КИБЕРФИЗИЧЕСКИЕ СИСТЕМЫ

<i>Деева И.Ю.</i> Многомасштабное моделирование цифрового образа человека в киберпространстве.....	5
<i>Корепанов П.Ю., Иващенко А.А.</i> Автоматическая генерация производственного плана.....	9
<i>Кормилицына Е. К.</i> Разработка средств балансирования вычислительной нагрузки в многопроцессорной системе.	13
<i>Гаврилов С.А., Кыздарбекова А.С., Резников С.С.</i> Нейрокомпьютерные интерфейсы для выявления сигналов управления бионическими устройствами на основе анализа физиологических и нервных процессов у человека.....	17
<i>Турсуков Н.О.</i> Модель функционирования группы беспилотных летающих аппаратов в среде симуляции v-гер	23
<i>Усова М.А., Вискнин И.И.</i> Модель цифрового производства.....	27

ТЕХНОЛОГИИ МУЛЬТИМЕДИА И КОМПЬЮТЕРНОЕ ЗРЕНИЕ

<i>Белова Д.А.</i> Применение технологий виртуальной реальности в психокоррекции	31
<i>Лемешев А.Ф.</i> Методы эффективной фильтрации синтезированного изображения для решения проблемы реалистичного рендеринга реального времени.....	35
<i>Меженин А.В., Трушина Ю.А.</i> Повышение качества запекания текстурных карт в процессе создания 3d контента	38
<i>Полякова Д.А., Перепелица Ф.А.</i> Особенности использования программных алгоритмов для оптимизации изделий в среде Autodesk Netfabb	44
<i>Смирнов А.А., Трунин Е.М., Исаев И.В., Чижиков И.В., Овсепьян А.Л., Бегиев О.Б., Мурт Л.Л., Кудлахмедов Ш.Ш., Назмиев А.И., Сизов П.А., Бабаев Д.Б.</i> Применение инс типа «персептрон» в морфологической диагностике папиллярного рака щитовидной железы.....	49
<i>Стоколяс И.А.</i> Подходы к автоматизации обработки изображения формата 360 градусов.....	57

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ. ПРОГРАММНАЯ ИНЖЕНЕРИЯ

<i>Аукен Г., Булозубов А.В.</i> Исследование применимости технологии блокчейн в области здравоохранения	61
---	----

<i>Деркач А.Н., Алексанков С.М.</i> Модель надежности кластера с поддержкой непрерывности вычислительного процесса при миграции виртуальных машин.....	65
<i>Орешин С.А., Лисицына Л.С.</i> Подход к прогнозированию поведения сложных систем на основе методов машинного обучения	69
<i>Матросова Н.Д., Штенников Д.Г.</i> Исследование зависимости результатов тестирования от времени изучения теоретического материала студентами на открытом наборе данных	73
<i>Миротавская Л.А., Бессмертный И.А.</i> Извлечение концептов из естественно-языковых текстов на основе контекста	77
<i>Нугманова А.А., Черных И.А., Ховричев, М.А., Булушева А.В.</i> Построение автоматических диалоговых систем для служб технической поддержки методами машинного обучения без учителя	81
<i>Пойманова Е.Д., Татарникова Т.М.</i> Управление пространственными ресурсами систем хранения данных.....	85
<i>Шульмина В.А.</i> «PLR»: программа проверки лабораторных работ по курсу «операционные системы».....	89

КОМПЬЮТЕРНАЯ БЕЗОПАСНОСТЬ

<i>Брыкина И.М., Маркина Т.А.</i> Уровни безопасности ИОТ	94
<i>Кулмаханов А.</i> Анализ методов защиты информации, передаваемой по волоконно-оптическим линиям связи.....	98
<i>Пакулин М.А.</i> Концепция сетевой безопасности в автоматизированных системах управления технологическим процессом (асу тп)	103
<i>Платонов Т.С. Оголюк А.А.</i> Межсетевые экраны уровня веб-приложения в современном мире.....	106
<i>Платонов Т.С.</i> Обзор характерных шаблонов работы шифровальщиков и способов их обнаружения.....	110

КИБЕРФИЗИЧЕСКИЕ СИСТЕМЫ

УДК 658.512

МНОГОМАСШТАБНОЕ МОДЕЛИРОВАНИЕ ЦИФРОВОГО ОБРАЗА ЧЕЛОВЕКА В КИБЕРПРОСТРАНСТВЕ

И.Ю. Деева

Университет ИТМО, 197101, Санкт-Петербург, Россия

Проведен аналитический обзор методов и технологий, обеспечивающих конструирование и использование цифровых образов объектов реального мира в различных прикладных областях. Целью работы является анализ и оценка современного состояния существующих методов и технологий, позволяющих моделировать цифровые образы объектов реального мира.

Ключевые слова: цифровой образ, цифровой след, цифровой двойник, машинное обучение, модель

Введение

В последнее время все более активно происходит процесс оцифровки объектов и процессов реального мира, т.е. создание их цифровых образов. Излишне говорить, что люди регулярно оставляют достаточно своих персональных данных в социальных сетях и на различных сайтах, формируя таким образом свой уникальный цифровой след. Тем не менее, до сих пор не существует единой концепции и методологии, которая позволила бы формировать универсальные цифровые модели личности в киберпространстве. Задача создания ЦО может быть решена на основе иерархической системы моделей в различных масштабах.

Обзор предметной области

Процесс разработки многомасштабной модели цифрового образа человека в киберпространстве включает в себя непосредственное построение этой модели, а также верификацию и валидацию полученного результата. Многомасштабная модель должна быть способна представить пользователя в разных измерениях, например, в соответствии с его личными данными (гражданство, возраст, пол, образование и т. д.), информацией в социальных сетях (фотографии, сообщения, количество друзей и т. д.) и ежедневной активностью на разных сайтах (клики, заказы в интернет-магазинах и т. д.).

Одним из способов построения многомасштабной модели ЦО может быть имитация поведенческой активности человека в социальных сетях. Для моделирования этого процесса как правило используются группы ботов, способных подражать поведению реального человека [1]. Такой подход к моделированию особенно полезен, когда требуется выявить основные отклонения в поведении пользователей. Например, [5] использует данный метод в задачах обнаружения спама. Также возможно построение композитных моделей в рамках ЦО, где могут быть использованы механизмы логического вывода и семантического анализа с использованием онтологий.

Для динамической идентификации моделей на данных могут применяться методы машинного обучения в своей общей постановке. Большой класс методов представляют методы для оценки параметров и выращивания членов регрессионных моделей. Для решения

этой задачи могут быть использованы ансамбли случайных деревьев (например, алгоритмы *adaboost*, *xgboost*) и нейронные сети (например, сети радиально-базисных функций и сети на основе обратного распространения ошибки) [2]. Для автоматизации процесса построения моделей на основе методов машинного обучения (например, нейронных сетей) могут использоваться подходы по выращиванию архитектуры и оптимизации гиперпараметров моделей [3, 4].

Однако для реализации всех указанных методов моделирования необходимо накопление критического объема исторической информации о пользователе, позволяющей создать эффективную и правдоподобную модель его поведения, действующую на протяженном отрезке времени от момента инициации, что осуществляется сейчас посредством повсеместного распространения цифровых коммуникаций и статистического анализа социальных сетей.

Выбор подхода к формированию модели может быть обоснован основными целями, которые ставит перед собой исследователь при формировании ЦО. Например, в [5] ЦО применяется в задачах изучения искажения информации.

Содержание исследования

Проведено исследование данных пользователей социальной сети Вконтакте. Набор данных представлял собой информацию о пользователях, которую они открыто публикуют на своих страницах. А именно:

- *id*;
- количество комментариев на странице;
- количество постов на странице;
- количество друзей;
- количество лайков на странице.

Посредством моделирования ЦО каждого пользователя предполагалось определить, как характеристики коррелируют с количеством друзей пользователя. Таким образом, целевой переменной являлось именно количество друзей. Такое моделирование прежде всего необходимо для дальнейшего предсказания количества друзей любого пользователя социальной сети.

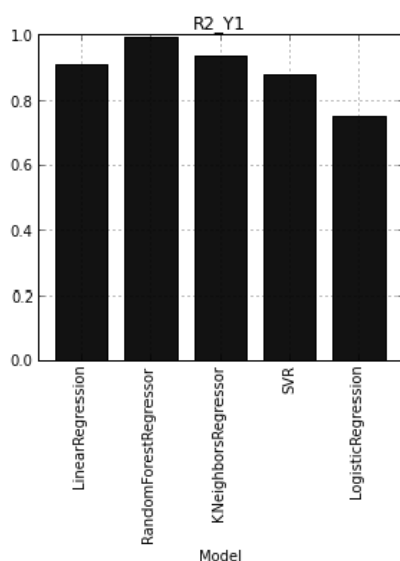


Рисунок 1. Результаты исследования Вконтакте

Для создания модели были использованы основные методы машинного обучения, в том числе метод наименьших квадратов, метод построения случайных лесов, метод К-

ближайших соседей, логистическую регрессию, а также метод опорных векторов с опорным ядром. В качестве оценки степени влияния каждого предиктора использовалась такая метрика, как коэффициент детерминации (R-квадрат). Результаты представлены на рисунке 1.

Из графика видно, что лучше всех с задачей справился метод случайных лесов, его коэффициент детерминации выше других. Такой результат можно объяснить тем, что исходные данные имели недостаточную интерпретируемость, такие сложные и разнородные данные как правило плохо аппроксимируются линейными моделями.

Также проведены были подобные исследования на данных из социальной сети Facebook. Здесь лучший результат показал метод опорных векторов, однако он также указал на обратную зависимость предикторов и отклика. Это можно объяснить тем, что возможно в социальной сети Facebook пользователю лайки чаще всего ставят не только друзья, но и сторонние пользователи. Результаты представлены на рисунке 2.

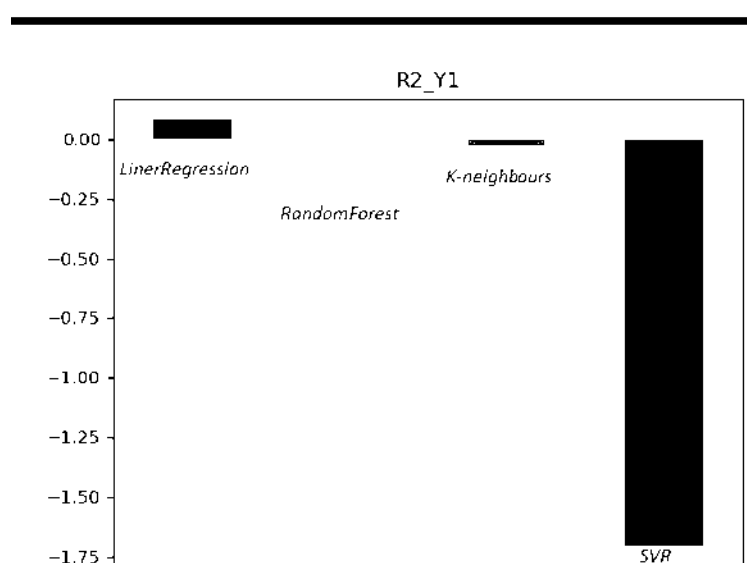


Рисунок 2. Результаты исследования Facebook

Заключение

В ходе аналитического обзора была продемонстрирована фрагментарность существующих подходов к созданию ЦО и ЦД, что требует активных исследований в области формирования единой концепции создания и использования ЦО личности в киберпространстве. Актуальность таких исследований также подкрепляется широким спектром возможностей использования цифровой модели личности, включающим прогнозирование поведения человека в киберпространстве, автоматизацию деятельности человека с замещением некоторых рутинных действий, создание бессмертной электронной копии человека, а также моделирование поведения целого социума.

ЛИТЕРАТУРА

1. Luo L. et al. Agent- based human behavior modeling for crowd simulation //Computer Animation and Virtual Worlds. – 2008. – Т. 19. – №. 3- 4. – С. 271-281.
2. Ahmadizar F. et al. Artificial neural network development by means of a novel combination of grammatical evolution and genetic algorithm //Engineering Applications of Artificial Intelligence. – 2015. – Т. 39. – С. 1-13.

3. Xue B. et al. A survey on evolutionary computation approaches to feature selection //IEEE Transactions on Evolutionary Computation. – 2016. – Т. 20. – №. 4. – С. 606-626.
4. Kovalchuk S.V. et al. A Conceptual Approach to Complex Model Management with Generalized Modelling Patterns and Evolutionary Identification //Complexity. – 2018.
5. Cheng Cao et al. Combating user misbehavior on social media // the Office of Graduate and Professional Studies of Texas A&M University. – 2017.

Сведения об авторе

Ирина Юрьевна Деева

— аспирант; Университет ИТМО,
Факультет информационных технологий и программирования;
E-mail: iriny.deeva@gmail.com

Ссылка для цитирования: Деева И.Ю. Многомасштабное моделирование цифрового образа человека в киберпространстве. // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 5—8.

MULTI-SCALE MODELING OF HUMAN DIGITAL ENTITY IN CYBERSPACE

I.Y. Deeva

ITMO University, 197101, St. Petersburg, Russia

In that article we present an analytical review of methods and technologies that provide for the design and use of digital entities of real-world objects in various applied areas. The aim of the work is the analysis and assessment of the current state of existing methods and technologies that allow to model digital entities of real-world objects.

Keywords: digital entity, digital track, digital twin, machine learning, model

Data on author

Irina Y. Deeva

— student; ITMO University,
Faculty of Information Technologies and Programming;
E-mail: iriny.deeva@gmail.com

For citation: Deeva I.Y. Multi-scale modeling of human digital entity in cyberspace. // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 5—8 (in Russian).

УДК 004.822

АВТОМАТИЧЕСКАЯ ГЕНЕРАЦИЯ ПРОИЗВОДСТВЕННОГО ПЛАНА

П.Ю. Корепанов, А.А. Иващенко*Университет ИТМО, 197101, Санкт-Петербург, Россия*

Исследованы подходы к моделированию предметной области по стандарту ISO 15926. Разработана архитектура программного обеспечения в рамках которой рассмотрена имитационная модель производства, позволяющая составлять план производственного процесса.

Ключевые слова: технологическая карта, ISO 15923, производственный процесс, операции, фонды предприятия

Введение

На сегодняшний день невозможно представить промышленные предприятия без интеллектуальных систем управления. В данной статье рассматривается разработка системы, принадлежащей к классу систем планирования производственного плана. Создание такой системы позволит избежать издержек и максимизировать прибыль. Данные проблемы являются актуальными, как в текущих реалиях промышленных гигантов, так и небольших производств. Разрабатываемые алгоритмы оптимизации позволят существенно увеличить эффективность производственных предприятий за счет сокращения простоев оборудования, увеличение загруженности персонала и оптимизации труда персонала.

Также не стоит забывать, что до сих пор расписание производств на многих предприятиях строится «вручную». При этом невозможно быстрое перераспределение ресурсов, что не позволяет производству оперативно реагировать на быстрое изменение ситуации и вносит ошибки, порождаемые человеческим фактором.

Таким образом, разрабатываемая система призвана частично решить перечисленные проблемы. В данной работе за основу взят классический подход к планированию на базе систем неравенств.

Новизна работы заключается в применении методов моделирования данных высших онтологий при формировании моделей предметной области [BORO, ISO 15926].

Обзор предметной области

Для того чтобы разрабатывать алгоритмы планирования в первую очередь необходимо формализовать предметную область. В данном случае формализуется сборочный цех со значимыми внутренними особенностями.

Как правило, на любом предприятии имеется специальный документ – технологическая карта, детально описывающая весь перечень операций, по достижению которых воспроизводится единица продукции. Технологическая карта хранит информацию о зависимостях между операциями, привязках ресурсов к операциям, трудоёмкостях операций, периодичность операций, результат каждой операции.

Далее важно учесть ресурсы предприятия. В рамках сборочного производства такими ресурсами могут быть: персонал, оборудование, организация конвейерной производственной линии.

Последним пунктом для построения модели производства является производственный план. То есть перечень продукции в составе заказа и срок реализации данного заказа. На рисунке 1 представлен пример взаимодействия перечисленных выше моделей.



Рисунок 1. Модель производства

Содержание исследования

На рисунке 2 представлена модель архитектуры программного обеспечения.

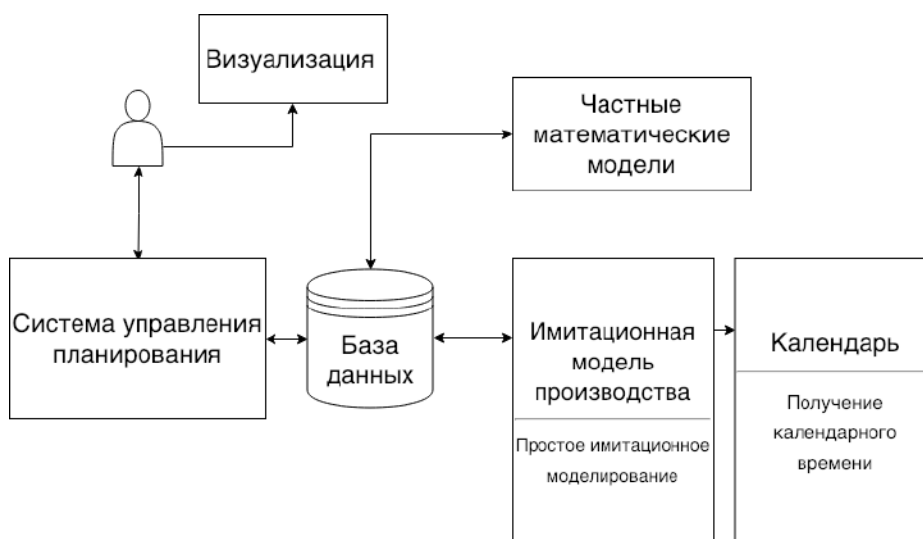


Рисунок 2. Архитектура ПО

Данная архитектура содержит следующие элементы:

- пользовательский интерфейс для взаимодействия с системой, который также позволяет получать информацию о работе системы в виде диаграмм, или графиков;
- система управления планирования взаимодействует со всеми элементами системы и является главным распорядителем задач;
- база данных хранит всю информацию о производстве и результаты планирования;
- имитационная модель производства создает план на основе информации из технологической карты и производственного плана;

- календарь обрабатывает абсолютные значения, используемые при планировании, и привязывает их к конкретным датам;
- частные оптимизационные модели работают с уже сформировавшимся планом, который получен в результате имитационного моделирования. К данному плану применяются алгоритмы оптимизации, зависящие от конкретных целей. Таким целями могут быть: задачи упорядочивания, задачи согласования, задачи распределения, задачи с суммарными критериями оптимизации, задачи с минимаксимальными критериями оптимизации [1, 2, 3].

Работа сосредоточена на разработке имитационной модели производства.

$$\begin{cases} t_{2A} = t_{1A} + 4 \\ t_{2B} = t_{1B} + 3 \\ t_{2C} = t_{1C} + 2 \\ t_{2A} \leq t_{1B} \\ t_{2A} \leq t_{1C} \end{cases}$$

$$\begin{cases} 0 \leq t_{1A} \\ 0 \leq t_{1B} \\ 0 \leq t_{1C} \end{cases} \begin{cases} 0 \leq 0 \\ 4 \leq t_{1B} \\ 4 \leq t_{1C} \end{cases} \begin{cases} 0 \leq 0 \\ 4 \leq 4 \\ 7 \leq t_{1C} \end{cases}$$

Рисунок 3. Система неравенств

Таким образом, планирование делится на шаги, каждый раз при этом формируются новые ограничения вводимые ресурсами [4].

Алгоритм планирования состоит из следующих этапов:

- Создание шаблона продуктов на основе заказа. Другими словами, воспроизведение порядка операций, указанного в технологической карте, статическая система неравенств (рис.3).
- Создание плана на основе заказа. На данном этапе алгоритм создает шаблоны всех продуктов, которые перечислены в заказе, также учитывая количество одноименной продукции.
- Пошаговая реализация плана с учетом ресурсных ограничений. Добавление дополнительных ограничений в систему неравенств, которые динамически меняются с каждым шагом планирования.

Заключение

В рамках работы были достигнуты следующие результаты:

1. Формализована предметная область, в результате создана база данных, которая хранит информацию о технологической карте, ресурсах предприятия, а также заказ.
2. Создана математическая модель производственных процессов для задач планирования, которая включает:
 - а. статическое планирование, то есть без учёта производственных ограничений;
 - б. динамическое планирование, с учетом ресурсных ограничений.
3. Произведено тестирование программы в различных режимах функционирования.

На основе всех вышеперечисленных пунктов разработана система, которая способна к планированию производственных процессов. В дальнейшем планируется создание частных

ОПТИМИЗАЦИОННЫХ моделей, которые позволят повысить эффективность целевой оптимизации.

ЛИТЕРАТУРА

1. *Pinedo, M.L.*, Planning and Scheduling in Manufacturing and Services: Second Edition / Planning and scheduling in manufacturing and services: Second edition, Springer, New York, pp. 1 – 536
2. *Gjeldum, N., Tufekcic, D., Veza, I.*, Modeling of control loop in production scheduling and inventory level control / Annals of DAAAM and Proceedings of the International DAAAM Symposium, 2007, pp. 297-298
3. *Hahn, Gerd J., Brandenburg, M.*, A sustainable aggregate production planning model for the chemical process industry / Computers & operations research, 2018, 94, pp. 154-168
4. *Theuer, H., Leo, A.K.*, Market survey | [IT in der Produktion]/ Productivity Management, 2011, 16(4), pp. 37-45

Сведения об авторе

- | | |
|---|---|
| Павел Юрьевич Корепанов | — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: p.corepanow1@yandex.ru |
| Аркадий Александрович
Иващенко | — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: ivaschenkoarkady@gmail.com |

Ссылка для цитирования: Корепанов П.Ю., Иващенко А.А. Автоматическая генерация производственного плана // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 9—12.

AUTOMATIC GENERATION OF FACTORY PRODUCTION SCHEDULE

P.Y. Korepanov, A.A. Ivashchenko

ITMO University, 197101, St. Petersburg, Russia

The approaches of modeling the subject area according to the ISO 15926 standard were explored. A software architecture had been developed within which a production simulation model was considered, which allows for the preparation of a production process plan.

Keywords: production routing , ISO 15923, production process, operations, production funds

Data on author

- | | |
|------------------------------|--|
| Pavel Y. Korepanov | — student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: p.corepanow1@yandex.ru |
| Arkady A. Ivashchenko | — student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: ivaschenkoarkady@gmail.com |

For citation: Korepanov P.Y., Ivashchenko A.A. Automatic generation of factory production schedule. // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 9—12 (in Russian).

УДК 004.451.26

РАЗРАБОТКА СРЕДСТВ БАЛАНСИРОВАНИЯ ВЫЧИСЛИТЕЛЬНОЙ НАГРУЗКИ В МНОГОПРОЦЕССОРНОЙ СИСТЕМЕ

Е.К. Кормилицына

Университет ИТМО, 197101, Санкт-Петербург, Россия

Объектом исследования являются многопроцессорные системы, предпочтение которым отдают все большее количество компаний по разработке встраиваемых систем. Для того, чтобы преимущества перехода к многопроцессорным системам были не только в теории, но и на практике, разработчику многопроцессорных систем придется решить проблему распределения вычислений по процессорам с целью достижения эффективной загрузки каждого, а также проблему преодоления коллизий в случаях, когда несколько процессоров пытаются получить доступ к одному ресурсу. В связи с чем целью работы является повышение производительности системы и снижение времени создания программного обеспечения за счет разработки автоматизированного механизма диспетчеризации задач по узлам системы, обеспечивающего балансировку вычислительной нагрузки. В работе проводится анализ готовых коммерческих и академических решений по балансировке загрузки процессоров в многопроцессорных системах с целью выявления и устранения их недостатков и синтеза их достоинств. Происходит выбор архитектуры многопроцессорной системы, так как от нее непосредственно зависит алгоритм планирования. Описывается разработка и функционирование алгоритма планирования задач в многопроцессорной системе, который направлен на обеспечение сбалансированности процессоров системы. В дополнение проводится анализ эффективности распределения задач, произведенного разработанным в проекте алгоритмом глобального планирования, по выбранным критериям.

Ключевые слова: многопроцессорные системы, встраиваемые системы, планирование, реальное время, балансировка, диспетчеризация

Введение

На сегодняшний день сильно возрастает потребность в повышении производительности вычислительных систем. Это обусловлено тем, что зачастую очень сложно, а порой даже невозможно из-за нехватки мощности одноядерного процессора выполнять обработку данных в реальном времени, кроме того появляется необходимость выполнять параллельную обработку данных. В области разработки встраиваемых систем данная проблема также актуальна.

Современные исследования направлены на:

- разработку и интегрирование более мощных одноядерных процессоров;
- использование многоядерных или многопроцессорных систем.

Все большее количество компаний отдает предпочтение многопроцессорным системам. Эффективное распределение задач и балансировка нагрузки процессоров – это важный для производительности многопроцессорных систем вопрос. За планирование задач в операционных системах реального времени отвечает планировщик задач. Следовательно, от

планировщика требуется такое распределение задач по процессорам системы, чтобы последние имели равную вычислительную нагрузку.

Обзор предметной области

В ходе работы был проведен анализ готовых коммерческих и академических решений. К коммерческим относятся следующие проекты: OCPB QNX Neutrino, OCPB RTEMS, которые поддерживают многопроцессорность встраиваемых систем, описание данных работ представлено в работах [1, 2]. К академическим относятся экспериментальная работа инженеров из Йоркского университета, а именно, адаптация FreeRTOS для работы в симметричной многопроцессорной системе, а также TOPPERS/FDMP Kernel – проект ученых из университета Нагоя, с данными проектами можно ознакомиться в работах [3, 4]. Не все архитектурные решения современных коммерческих и академических предложений позволяют масштабировать систему, что должно способствовать повышению производительности и скорости реакции. Также рассматриваемые решения не способствуют повышению производительности вследствие использования разделяемой памяти и топологии «шина» для связи процессоров, которые накладывают дополнительные условия на совместный доступ к памяти и синхронизацию кэшей.

Содержание исследования

На сложность реализации многопроцессорных систем и эффективность использования ресурсов влияет выбор архитектуры системы. В ходе работы была предложена архитектура, позволяющая разрабатывать программный код на однопроцессорной системе и в дальнейшем переносить его на массив физически разделенных процессоров. Таким образом, разработанную архитектуру можно назвать основанной на кластерном подходе, в котором регламентировано закрепление задач за конкретным процессором. Схематично архитектура представлена на рисунке 1.

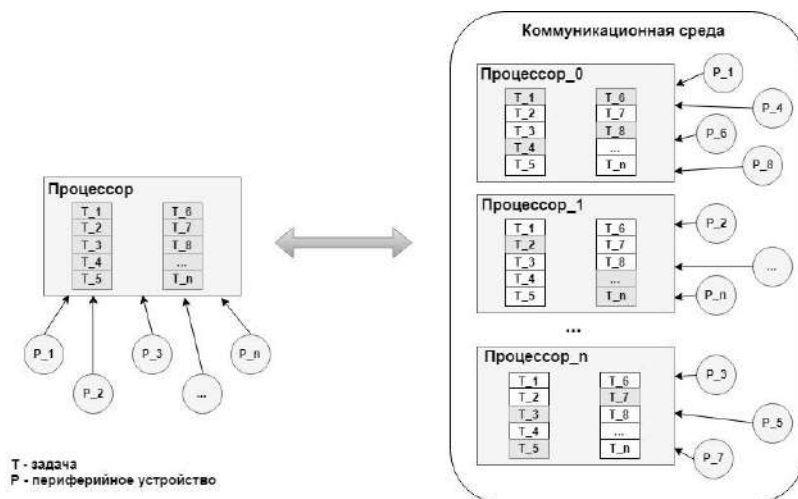


Рисунок 1. Архитектура многопроцессорной системы

Планирование выполнения задач осуществляет глобальный статический планировщик, который был разработан в рамках данной работы. На рисунке 2 представлен алгоритм работы глобального планировщика в упрощенном виде. Основная задача глобального планировщика – это балансировка нагрузки системы, для решения которой ему необходимо иметь сведения о вычислительных ядрах, задачах и их взаимодействии, подробнее с принципами работы глобальных планировщиков можно ознакомиться в работе [5]. В результате работы глобального планировщика задачи распределены между существующими вычислительными узлами так, что ресурсы системы используются оптимально и

рационально, а время обслуживания событий сокращено. Для оценки функционирования глобального планировщика и системы в целом, а также обеспечения эффективного обслуживания заявок в ней глобальный планировщик был построен на основе модели массового обслуживания и использовались следующие характеристики: коэффициент загрузки системы, коэффициент простоя системы, вероятность потери заявки, расчет среднеквадратического отклонения, формулы данных критериев представлены в работе [6].

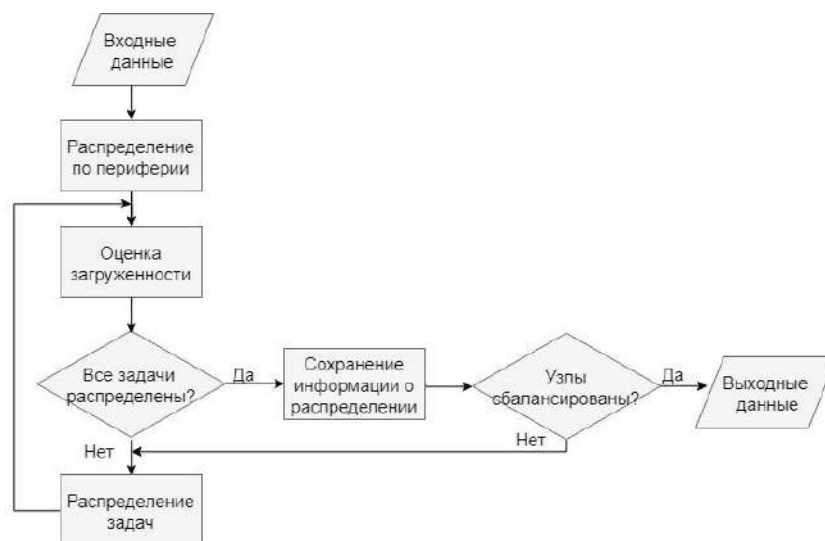


Рисунок 2. Упрощенная схема работы глобального планировщика

Заключение

В работе была выбрана и описана архитектура многопроцессорной системы, которая компенсирует недостатки текущих архитектур. А именно, позволяет систему масштабировать и повысить производительность. Была выбрана топология связи вычислительных узлов и разработан протокол коммутации. Они направлены на достижение максимально полезной загрузки узлов с минимальной передачей избыточной информации, которая влияет на время передачи. Предложен глобальный статический алгоритм планирования задач в многопроцессорной системе, направленный на обеспечение балансировки загрузки системы. Статическое планирование позволяет повысить предсказуемость системы, что важно для систем реального времени. Работа предложенного алгоритма планирования была оценена с помощью вероятностных и экономических характеристик: коэффициент загрузки системы, коэффициент простоя системы, вероятность потери задач. Опираясь на данные характеристики, распределение задач можно проводить неоднократно, добиваясь эффективного. Для оценки сбалансированности вычислительных узлов использовалось среднеквадратическое отклонение.

ЛИТЕРАТУРА

1. QNX Neutrino RTOS System Architecture. - QNX Software Systems International Corporation. – 2006. – 411 p.
2. RTEMS Classic Guide. - The RTEMS Documentation Project. – 2016. – 716 p.
3. Mistry J., Naylor M. Adapting FreeRTOS for Multicore. - Wiley InterScience. – 2013. – 27 p.
4. Tomiyama H., Honda S. Real-Time Operating Systems for Multicore Embedded Systems. - Graduate School of Information Science Nagoya University. – 2008. – 6 p.

5. *Liberato F., Lauzac S., Melhem M., Mosse D.* Fault tolerant real-time global scheduling on multiprocessors. – Real time systems. Proceedings of the 11th Euromicro Conference on. – 2012. – 9 p.
6. *Алиев Т.И.* Основы моделирования дискретных систем. – СПбГУ ИТМО. – 2009. – 363с.

Сведения об авторе

Елизавета Константиновна Кормилицына — аспирант; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: sholohova.elizaveta@mail.ru

Ссылка для цитирования: *Кормилицына Е.К.* Разработка средств балансирования вычислительной нагрузки в многопроцессорной системе. // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 13—16.

DEVELOPMENT OF TOOLS FOR BALANCING THE COMPUTATION LOAD IN MULTIPROCESSOR SYSTEM

Е.К. Kormilitsyna

ITMO University, 197101, St. Petersburg, Russia

The aim of the research is to increase productivity and reduce the software development time for multiprocessor real time system. It can be achieved by developing tools of automatic tasks scheduler between processors, which provides balanced load for all processors in the entire system. The modern commercial and academic solutions for processors load balancing were analyzed and their advantages and disadvantages have been identified. Also was chosen architecture for multiprocessors system and developed the tool that provides RTOS tasks scheduling and optimal processors load balancing algorithm in multiprocessors system.

Keywords: multiprocessor, real time, symmetric multiprocessing architecture, asymmetric multiprocessing architecture, scheduling

Data on author

Elizaveta K. Kormilitsyna — post-graduate student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: sholohova.elizaveta@mail.ru

For citation: *Kormilitsyna E.K.* Development of tools for balancing the computation load in multiprocessor system // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 13—16 (in Russian).

УДК 004.89

НЕЙРОКОМПЬЮТЕРНЫЕ ИНТЕРФЕЙСЫ ДЛЯ ВЫЯВЛЕНИЯ СИГНАЛОВ УПРАВЛЕНИЯ БИОНИЧЕСКИМИ УСТРОЙСТВАМИ НА ОСНОВЕ АНАЛИЗА ФИЗИОЛОГИЧЕСКИХ И НЕРВНЫХ ПРОЦЕССОВ У ЧЕЛОВЕКА

С.А. Гаврилов, А.С. Кыздарбекова, С.С. Резников

Университет ИТМО, 197101, Санкт-Петербург, Россия

В настоящее время, во всем мире активно ведутся разработки различных нейрокомпьютерных интерфейсов (НКИ) для управления протезами, экзоскелетами, роботами, для взаимодействия с виртуальной реальностью, для реабилитационной медицины и даже для профессионального спорта. Особенно хотелось бы выделить реабилитационную медицину, и протезирование конечностей человека. Разработки и исследования в данном направлении помогут частично, а в перспективе и полностью восстановить функциональные возможности людей с ампутацией или врожденными дефектами конечностей. Таким образом, целью данной работы является обзор методов регистрации сигналов управления бионическими устройствами на основе анализа физиологических и нервных процессов у человека. В связи с этим, в первую очередь, интересна разработка датчика и контроллера для дальнейшего применения в системах роботизированных протезов. На данный момент изготовлен первый прототип с пятью каналами для подключения датчиков и интерфейсом USB для связи с персональным компьютером. Также изготовлен первый прототип с 5-ю каналами для подключения датчиков и интерфейсом USB для связи с персональным компьютером.

Ключевые слова: нейрокомпьютерный интерфейс, протезирование, электромиография

Введение

В настоящее время, во всем мире активно ведутся разработки различных нейрокомпьютерных интерфейсов (НКИ) для управления протезами, экзоскелетами, роботами, для взаимодействия с виртуальной реальностью, а также реабилитационной медицине и даже в спорте [1]. Кампании, такие как Touch Bionics (Великобритания), Biometrics Ltd (Великобритания), Delsys (США), Noraxon (США), Ottobock (Германия), Mind Technologies (США), Covidien (Ирландия) и многие другие не первый год занимаются коммерческими разработками в этой области. Даже Илон Маск приобрёл и возглавил Neuralink. Из российского сегмента можно бы выделить компании: NeuroG, iBrain, «Объединенная приборостроительная корпорация», Cyber Myonics и Моторика.

Особенно хотелось бы выделить направление реабилитационной медицины, и в частности протезирование конечностей человека [2]. Разработки и исследования в данном направлении помогут частично, а в перспективе и полностью восстановить функциональные возможности людей с ампутацией или врожденными дефектами конечностей.

Обзор предметной области

Можно выделить два основных вида НКИ для регистрации управляющих сигналов: регистрация и обработка данных об активности мозга регистрируемых по средствам электроэнцефалографии (ЭЭГ) и регистрация, обработка биопотенциалов мышц по

средствам электромиографии (ЭМГ). В свою очередь электромиография бывает инвазивной и неинвазивной (поверхностной). Также хотелось бы выделить в отдельный тип интерфейсы нейропротезов и нейроимплантов разрабатываемые в рамках нейропротезирования (сенсорного протезирования). Самый распространённый – это кохлеарный имплантат, предназначенный для компенсации потери слуха при нейросенсорной (перцептивной) тугоухости. Есть принципиальные отличия между интерфейсами нейроимплантов и НКИ в направлении передачи информации. Нейроимпланты осуществляют возбуждение нервных окончаний, относящихся к сенсорной системе и реализуют функцию передачи из внешней среды в нервную систему. НКИ предназначен для передачи информации из центральной нервной системы в электронно-вычислительный комплекс.

К перспективным технологиям для реализации НКИ можно отнести магнитоэнцефалографию. В Марте 2018 года учёными из Ноттингемского университета и Университетского колледжа Лондона была опубликована статья в журнале Nature о магнитоэнцефалографическом сканере мозга в виде шлема, позволяющего снимать показания в движении [3]. Технология на данный момент имеет ряд ограничений для реализации НКИ в связи с техническими особенностями. По моему мнению, есть определенные перспективы для реализации подобного интерфейса.

Нейрокомпьютерный Интерфейс На Основе ЭЭГ (ЭЭГ НКИ)

Данный НКИ основана на регистрации электрической активности мозга по средствам измерения электрических потенциалов поверхностными электродами, установленными в определенных точках на поверхности кожи головы и последующем математическом анализе полученных данных.

Механизм реализации достаточно сложен и состоит из нескольких этапов. В первую очередь формируется набор ЭЭГ-паттернов, соответствующих определенной метальной активности, воображаемому действию или вызванному потенциалу. На основе полученного набора производится обучение классификатора и последующее распознавание повторного появления паттерна. Набор ЭЭГ-паттернов для каждого человека индивидуален [4].

Данный метод универсален и его возможно применять независимо от характера ампутации или дефекта, а также, он более эффективен для взаимодействия с виртуальной реальностью. Но в тоже время, в связи со сложным математическим аппаратом и большим объёмом обрабатываемых данных, требуются значительные вычислительные мощности. Электроэнцефалография (ЭЭГ) – это неинвазивный метод исследования функционального состояния головного мозга [5]. Инвазивный метод функционального исследования коры полушарий головного мозга - электрокортикография.

Нейрокомпьютерный интерфейс на основе поверхностной ЭМГ (пЭМГ НКИ)

Данный интерфейс основан на регистрации и обработке биоэлектрических потенциалов в скелетных мышцах человека по средствам установки накладных электродов на поверхность кожи непосредственно над измеряемой мышцей и последующем математическом анализе полученных данных.

Механизм реализации значительно проще, чем у ЭЭГ НКИ и заключается в непосредственном измерении биоэлектрической активности мышцы, и преобразовании измерений в управляющее воздействие исполнительного механизма роботизированного протеза [6].

Основными сложностями реализации данного НКИ являются: «Плавающее» электрическое сопротивление поверхности кожи (аналогичная проблема возникает и в случае ЭЭГ НКИ); Перекрестные помехи или «Затекание» биоэлектрического сигнала - посторонний сигнал нейромускульной активности от других мышц, распространяемый по поверхности кожи [7]. Данный интерфейс обеспечивается простым математическим аппаратом в следствии чего, анализ данных и формирование управляющего воздействия

могут быть реализованы на базе встраиваемых систем, обладающих малым размером и низким энергопотреблением, а они в свою очередь являются важными факторами для протезирования.

Самым интересным проектом пЭМГ НКИ является браслет «Муо» от компании «Thalnic Labs». Коммерческая реализация уже запущена, а применение данного НКИ для управления роботизированным протезом активно развивается специалистами частного исследовательского университета Джонса Хопкинса (США) [8].

Специалисты Массачусетского технологического института в апреле 2018 года представили новый НКИ «AlterEgo» основанный на принципе регистрации субвокализации по средствам пЭМГ лицевых мышц [9]. Данная разработка значительно расширяет возможности применения пЭМГ НКИ не только для протезирования конечностей, но и для взаимодействия с виртуальной реальностью.

Нейрокомпьютерный интерфейс на основе инвазивной ЭМГ (ЭМГ НКИ)

ЭМГ НКИ аналогичен пЭМГ НКИ за исключением метода подключения. Инвазивное подключение электродов позволяет перейти от глобальной электромиограммы мышцы к электромиограмме отдельного мышечного волокна что дает ряд преимуществ перед пЭМГ НКИ:

- Исключение электрического сопротивления поверхности кожи.
- Увеличение отношения сигнал/шум.
- Значительное снижение перекрестных помех.
- Регистрация нейромускульной активности отдельного волокна.

Содержание исследования

В первую очередь, была поставлена задача по разработке датчика биопотенциала мышц. На рис.1 показаны прототипы датчиков мышечных биопотенциалов.

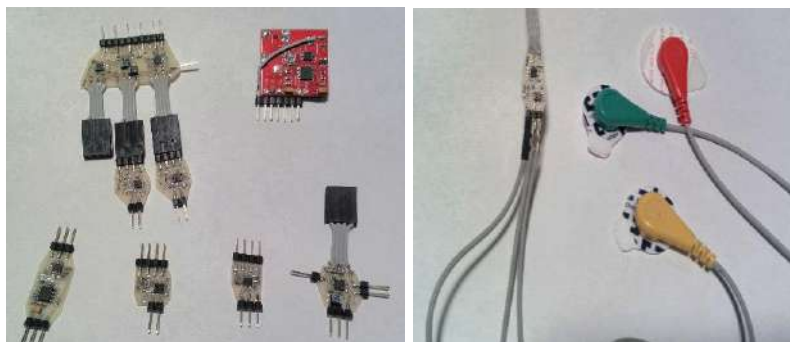


Рисунок 1. Прототипы датчиков мышечных биопотенциалов

Разработаны несколько аппаратных решений датчика и получены первые положительные результаты их работы.

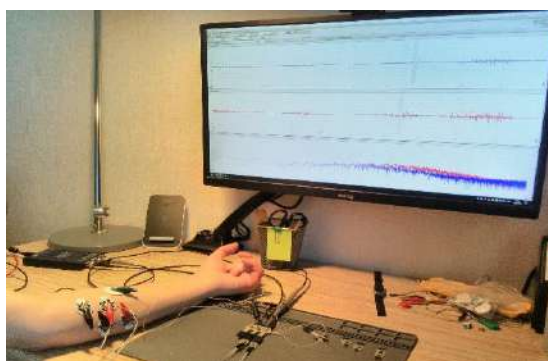


Рисунок 2. Испытание прототипов датчиков мышечных биопотенциалов

Параллельно ведется разработка контроллера НКИ. На данный момент изготовлен первый прототип с 5-ю каналами для подключения датчиков и интерфейсом USB для связи с персональным компьютером.

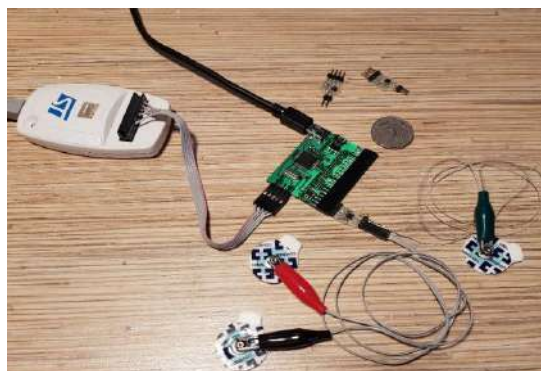


Рисунок 3. Прототип контроллера НКИ

Исходя из выше сказанного, пЭМГ НКИ является самым доступным и функциональным нейрокомпьютерным интерфейсом для разработки по ряду причин:

- Доступность и низкая стоимость элементной базы;
- Отсутствие необходимости непосредственного участия медицинских специалистов;
- пЭМГ НКИ более функционален для применения в системах роботизированных протезов конечностей;
- Имеется опыт разработки микроэлектроники и программного обеспечения для встраиваемых систем.

Заключение

Исходя из выше сказанного, пЭМГ НКИ является самым доступным и функциональным нейрокомпьютерным интерфейсом для разработки. Преимуществом является: доступность и низкая стоимость элементной базы, отсутствие необходимости непосредственного участия медицинских специалистов, также пЭМГ НКИ более функционален для применения в системах роботизированных протезов конечностей. Существующие разработки в Российской Федерации базируются на иностранных аппаратных решениях. В связи с этим, в первую очередь, интересна разработка собственного датчика и контроллера для дальнейшего применения в системах роботизированных протезов.

На данный момент авторами разработаны несколько аппаратных решений датчика биопотенциала мышц и получены первые положительные результаты их работы. А также параллельно ведутся разработки контроллера НКИ. Изготовлен первый прототип с 5-ю каналами для подключения датчиков и интерфейсом USB для связи с персональным компьютером. Далее, планируется проведение испытаний и оптимизация схемы прототипа контроллера, оптимизация схемы и проработка «сухого» подключения электродов датчика, а также разработка программного обеспечения для передачи данных в реальном времени с контроллера НКИ в пакет приложений для системного анализа (MATLAB, Scilab и др.) для дальнейшего исследования.

ЛИТЕРАТУРА

1. Шишкин А.В., Митин А.Е., Филиппова С.О. Проблема применения электромиографии с целью повышения эффективности тренировочного и соревновательного процессов в адаптивном. Современные проблемы науки и образования 6, 57-63 (2013)

2. Rubana H. Chowdhury, Mamun B.I. Reaz, Mohd Alauddin Bin Mohd Ali, Ashrif A.A. Bakar, Chellappan K.: Chang Surface Electromyography Signal Processing and Classification Techniques 13, 12431-12466 (2013).
3. Boto E., Holmes N., Leggett J., Roberts G., Shah V., Meyer S., Duque M., Mullinger J., Tierney M., Bestmann S., Barnes R., Bowtell R. & Matthew J.: Moving magnetoencephalography towards real-world applications with a wearable system. *Brookes in Nature* 555, 657–661 (2018)
4. Мокиенко О.А., Люкманов Р.К., Черникова Л.А., Супанова Н.А., Пирадов М.А., Фралов А.А. :Интерфейс мозг-компьютер: первый опыт клинического применения в России. *Физиология человека* 42 (1), 31-39 (2016)
5. Sadikoglu F., Kavalcioglu C., Dagman B.: Electromyogram (EMG) signal detection, classification of EMG signals and diagnosis of neuropathy muscle disease. *Procedia Computer Science*, 422-429 (2017)
6. Makhrov S.S., Erokhin S.D.: Analysis of methods for the formation and isolation of EEG patterns when registering signals in neural interfaces. *Telecommunications and transport* 11 (10), 55-59 (2017)
7. Рукина Н.Н., Кузнецов А.Н., Борзиков В.В., Комкова О.В., Белова А.Н. Поверхностная электромиография: ее роль и потенциал в развитии экзоскелета (обзор). *Современные технологии в медицине* 8 (2), 109–118 (2016)
8. Official website, // Thalmic Labs URL: <https://www.thalmic.com/>
9. Hardesty L., URL: <http://news.mit.edu/2018/computer-system-transcribes-words-users-speak-silently-0404>

Сведения об авторе

Степан Александрович Гаврилов	— аспирант; Университет ИТМО, Факультет систем управления и робототехники; E-mail: itgavrilov@corp.ifmo.ru
Айдана Садвакасовна Кыздарбекова	— аспирант; Университет ИТМО, Факультет систем управления и робототехники; E-mail: dok_k-shu@mail.ru
Станислав Сергеевич Резников	— доцент; Университет ИТМО, Факультет систем управления и робототехники; E-mail: sreznikov@corp.ifmo.ru

Ссылка для цитирования: Гаврилов С.А., Кыздарбекова А.С., Резников С.С. Нейрокомпьютерные интерфейсы для выявления сигналов управления бионическими устройствами на основе анализа физиологических и нервных процессов у человека // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 17—22.

BRAIN-COMPUTER INTERFACES FOR THE DETECTION OF SIGNALS CONTROLLING A BIONIC DEVICES BASED ON THE ANALYSIS OF THE PHYSIOLOGICAL AND NEURAL PROCESSES IN HUMANS**S.A. Gavrilov, A.S. Kyzdarbekova, S.S. Reznikov***ITMO University, 197101, St. Petersburg, Russia*

Currently, around the world are actively developing various neurocomputer interfaces (NICS) to control prostheses, exoskeletons, robots, to interact with virtual reality, for rehabilitation medicine and even for professional sports. I would especially like to highlight rehabilitation medicine, and prosthetics of human limbs. Development and research in this area will help partially, and in the long term and fully restore the functionality of people with amputation or congenital defects of the limbs. Thus, the purpose of this work is to review the methods of registration of control signals of bionic devices based on the analysis of physiological and nervous processes in humans. In this regard, first of all, it is interesting to develop a sensor and controller for further use in robotic prosthesis systems. At the moment, the first prototype with five channels for connecting sensors and a USB interface for communication with a personal computer has been made. Also made the first prototype with 5 channels for connecting sensors and USB interface for communication with a personal computer.

Keywords: brain-computer interface, prosthesis, electromyography

Data on author

- | | | |
|---------------------------------------|---|--|
| <i>Stepan. A. Gavrilov</i> | — | post-graduate student; ITMO University,
Faculty of Control Systems and Robotics;
E-mail: itgavrilov@corp.ifmo.ru |
| <i>Aidana. S. Kyzdarbekova</i> | — | post-graduate student; ITMO University,
Faculty of Control Systems and Robotics;
E-mail: dok_k-shu@mail.ru |
| <i>Stanislav S. Reznikov</i> | — | docent; ITMO University,
Faculty of Control Systems and Robotics;
E-mail: sreznikov@corp.ifmo.ru |

For citation: *Gavrilov S.A., Kyzdarbekova A.S., Reznikov S.S.* Brain-computer interfaces for the detection of signals controlling a bionic devices based on the analysis of the physiological and neural processes in humans // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 17—22 (in Russian).

УДК 004.942

МОДЕЛЬ ФУНКЦИОНИРОВАНИЯ ГРУППЫ БЕСПИЛОТНЫХ ЛЕТАЮЩИХ АППАРАТОВ В СРЕДЕ СИМУЛЯЦИИ V-REP

Н.О. Турсуков

Университет ИТМО, 197101, Санкт-Петербург, Россия

В работе реализована модель взаимодействия беспилотных летающих аппаратов с помощью фреймворка V-REP, а также проведен ряд экспериментов, с целью проанализировать возможные деструктивные воздействия на систему. В ходе разработки были сформированы данные о потребляемой аппаратами энергии, выявлены и изучены уязвимые элементы работы системы. Проведенные эксперименты позволили выявить необходимость повышения защиты от деструктивного информационного воздействия.

Ключевые слова: симуляция поведения роботов, беспилотные летающие аппараты, деструктивное воздействие, децентрализованная система, взаимодействие роботов

Введение

При проектировании сложных систем с участием роботов, подразумевающих их взаимодействие друг с другом, необходимо анализировать как скорость выполнения поставленных целей, так и наличие возможных непредсказуемых ситуаций. Особенно это стоит учитывать, если роботы (агенты) не зависят от централизованной системы, контролирующей их. Непредсказуемые ситуации могут возникать в случае преднамеренного или непреднамеренного информационного воздействия (ИВ) на агента [1, 4]. В подобных случаях целесообразнее всего проконтролировать поведение роботов в симуляторе, а также провести эксперименты с возможными вариантами деструктивного воздействия, с целью проанализировать возможные потери.

Обзор предметной области

В работе реализована модель функционирования группы беспилотных летающих аппаратов (БПЛА) при помощи фреймворка для симуляции роботизированных систем V-REP. Фреймворк представляет собой среду для построения и симуляции поведения роботов, без необходимости физического доступа к аппарату [2]. Особенности данного программного обеспечения в том, что встроенные плагины, описывающие работу БПЛА, можно изменять, а также добавлять собственные скрипты, для реализации поведения роботов. Программирование реализовано в симуляторе, с помощью редактора, поддерживающего скриптовый язык программирования Lua. Симулятор поддерживает визуализацию действий аппаратов в трехмерном виде, что еще более наглядно покажет их взаимодействия [3].

Содержание исследования

Основные характеристики идеального случая функционирования системы:

- квадратная область (далее – Помещение), по которой перемещаются агенты;
- 10 флагов - 10 задач;
- потрачено минимум времени на выполнение задачи;
- выполнены все задачи;

— энергия робота – 50 у.е

Помещение, по которому перемещаются агенты, представлено на рисунке 1. Она разделена на 100 квадратов, 10 из которых заняты под флаги, которые собирают агенты. Точка сдачи флага является также точкой старта (А). Каждый робот обладает достаточной энергией, для выполнения 2 задач.

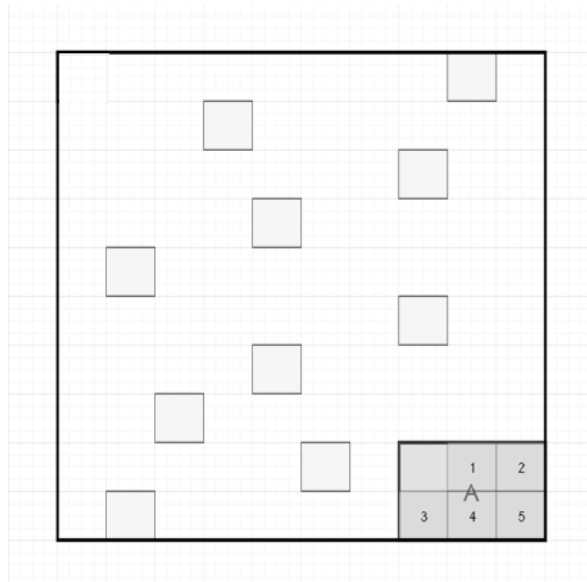


Рисунок 1. Схема Помещения

Ход работы системы:

Группа агентов начинает функционировать в области А на территории Помещения. Целью группы является выполнение всех задач. Задачей является перемещение флага в область А. Окончанием эксперимента является достижение цели.

Этапы, которые были выполнены для реализации эксперимента:

- создание группы агентов;
- введение алгоритма распределения задач между агентами;
- реализация выполнения задач агентами;
- организация деструктивного воздействия на реализованную модель;
- сбор полученных данных о времени на реализацию и эффективности выполнения задач и их последующее сравнение.

Определим основные виды деструктивного информационного воздействия, которому могут быть подвержены агенты:

- Агент не участвует в ИВ, но имеет такую возможность.
 - Косвенный ущерб – издержки агентов увеличиваются.
- Агент врет о своем местоположении, следовательно, выполняет не приоритетную для него цель.
 - Прямой ущерб – издержки агентов могут увеличиться, что приведет к уменьшению количества выполненных задач.
- Агент забирает задачу, но не выполняет ее.
 - Прямой ущерб – забранная агентом задача не выполняется

Движение и взаимодействие агентов реализовано во встроенном редакторе фреймворка. При начале работы, а также при посадке на площадку, агент движется по оси Z. Расчет выбора ближайшей цели реализован на основе координат флагов, которые знает каждый агент. В ходе работы системы, самый первый агент выбирает ближайшую цель, сообщает остальным агентам о своей задаче и отправляется на ее выполнение [5]. После выполнения задачи агент проверяет, есть ли еще не взятые флаги, и, если у него хватает

энергии, повторяет вышеописанный алгоритм. Информация о потраченной энергии и выполненных целях заносится в отдельный файл, информация из которого впоследствии и будет анализироваться. Принцип работы диверсантов в проведенных экспериментах, не отличается от работы остальных агентов, за исключением того, что информация, переданная остальным участникам системы (о выполнении задачи, о приоритетности задачи) зачастую неверная.

Все эксперименты проводились с одинаковым набором флагов, расставленных на поле и показанных на рисунке 2.

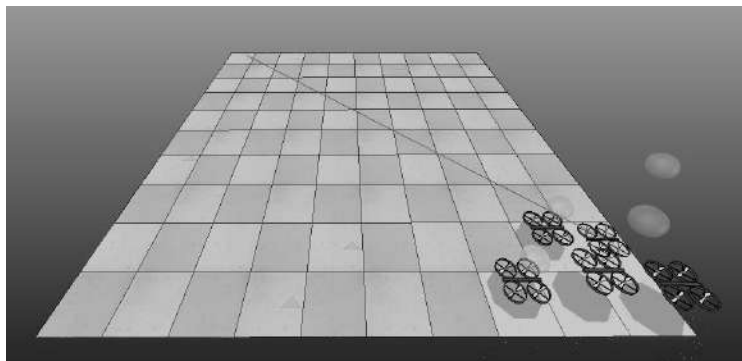


Рисунок 2. Модель работы БПЛА (диверсант взял дальнюю задачу)

В ходе проведения экспериментов был выявлен как прямой, так и косвенный ущерб. Наибольший ущерб системе наносят диверсанты, не выполняющие задач. Если диверсант не выполняет одну задачу, то, с большой вероятностью, один из агентов не сможет завершить работу из-за нехватки ресурсов. Схожий результат при условии, когда агент не выполняет ни одной задачи. Для решения подобных задач стоит либо увеличивать количество ресурсов у агентов, что будет ресурсозатратным, либо распределять движение агентов таким образом, чтобы агент с наибольшим количеством энергии выполнял лишние задачи, оставшиеся из-за диверсанта. В таком случае скорость выполнения цели значительно уменьшится.

Заключение

В ходе работы была реализована гибкая модель функционирования БПЛА, позволяющая анализировать и сравнивать потери при различных деструктивных воздействиях на работу системы, а также оценить поведение роботов в их обычном режиме работы.

ЛИТЕРАТУРА

1. Мордвинов Д.А., Литвинов Ю.В. Обзор применения формальных методов в робототехнике // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2016. №1 (236).
2. Сорокин С.В., Сорокина И.В., Солдатенко И.С. Использование виртуальных лабораторий в инженерном образовании // Инженерное образование, 2017.
3. Eric Rohmer, Surya P. N. Singh, Marc Freese, V-REP: A versatile and scalable robot simulation framework // 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems
4. Viksnin I.I. et al. Approaches to communication organization within cyber-physical systems // Open Innovations Association (FRUCT), 2017 20th Conference of IEEE. 2017. P. 484 – 490.
5. Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов. М.: ФИЗМАТЛИТ, 2009. 280 с

Никита Олегович Турсуков — *Сведения об авторе*
студент; Университет ИТМО,
Факультет безопасности информационных технологий;
E-mail:stepingnik@gmail.com

Ссылка для цитирования: Турсуков Н.О. Модель функционирования группы беспилотных летающих аппаратов в среде симуляции v-rep // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 23—26.

**MODEL OF FUNCTIONING OF UNMANNED AERIAL
VEHICLES IN THE ROBOT SIMULATION ENVIRONMENT V-REP**

N.O. Tursukov

ITMO University, 197101, St. Petersburg, Russia

The paper implements a model of interaction of unmanned aerial vehicles using the V-REP framework, as well as a number of experiments to analyze the possible destructive effects on the system. During the development, data on the energy consumed by the devices were formed, vulnerable elements of the system were identified. The experiments helped to identify the need for enhanced protection from the destructive informational influence.

Keywords: simulation of robot behavior, unmanned aerial vehicles, destructive impact, decentralized system, robot interaction

Nikita O. Tursukov — *Data on author*
student, ITMO University,
Faculty of Information Security and Computer Technologies;
E-mail:stepingnik@gmail.com

For citation: Tursukov N.O. Model of functioning of unmanned aerial vehicles in the robot simulation environment v-rep // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 23—26 (in Russian).

УДК 004.031.2

МОДЕЛЬ ЦИФРОВОГО ПРОИЗВОДСТВА

М.А. Усова, И.И. Вискнин*Университет ИТМО, 197101, Санкт-Петербург, Россия*

Предметом исследования является модель автономного цифрового производства. В работе описывается математическая модель автономного производства с использованием теоретико-множественного аппарата и понятий теории множеств. Рассматривается понятие информационного пространства в рамках киберфизической системы. Результаты работы могут быть использованы как основа для анализа защищенности системы с точки зрения информационной безопасности.

Ключевые слова: цифровое производство, индустрия 4.0, киберфизические системы

Введение

К концу XX века в производственной сфере использовались различные подходы к созданию автоматизированных производств, снижающих долю участия человека в технологическом производстве товаров и услуг. С возрастанием потребностей общества возникла необходимость в пересмотре подходов к организации производства и переходу к высокотехнологичным, “умным”, цифровым производствам, являющимся полностью автономными. Концепцию таких производств называют Индустрией 4.0. Основу ее составляют сенсоры, осуществляющие сбор информации, сеть Интернет в целом и облачные сервисы в частном. В рамках данной работы ставится задача создания модели цифрового производства и описание пространства информации, передаваемой внутри производства агентами-роботами.

Обзор предметной области

Концепция цифрового производства («умной фабрики») основана на взаимодействии элементов в рамках кибер-физической системы. Умная фабрика сочетает в себе методы Интернета вещей (IoT) и Интернета сервисов (IoS), которые обеспечивают взаимодействие физического уровня системы с информационным во время функционирования, позволяют реализовать бизнес-процессы на различных стадиях производства [1]. Подсистема Интернета вещей решает проблему на уровне взаимодействия и подключения роботов, Интернет сервисов адаптирует систему под изменения рынка товара, отслеживает спрос и предложение, отвечает за поддержку принятия решений и ценовую политику, отвечает за регулирование объема производства. Проблемы реализации концепции Индустрии 4.0 в области мануфактуринга не ограничиваются организацией функционирования системы, к ним также относят отсутствие законодательных актов, регулирующих их деятельность, стандартизацию систем, создание комплексной инфраструктуры для данной отрасли, обеспечение экологической и информационной, оценка эффективности использования ресурсов и функционирования логистической системы [2].

Авторы ряда работ [3-5] занимались исследованием вопроса организации функционирования группы автономных автомобилей. Методы данных исследований можно использовать для разработки логистической системы производства. В основе систем автономных автомобилей зачастую используются методы «подключенных машин» (connected cars), которые взаимодействуют друг с другом по каналу «Vehicle-to-vehicle» и с

объектами инфраструктуры по каналу «Vehicle-to-infrastructure». Такие системы позволяют координировать автономные транспортные средства, реализующие доставку товара от места его производства до потребителя.

Главный недостаток современных исследований в сфере Индустрии 4.0 заключается в отсутствии описательных и математических моделей производства, что не позволяет современным ученым изучить эффективность и результативность данной концепции предприятий.

Содержание исследования

Умное производство (умная фабрика) рассматривается как совокупность двух подсистем: производственная подсистема и логистическая подсистема:

$$System_{manufacturing} = Subsystem_{factory} \cup Subsystem_{Logistic}$$

Подсистема фабрики представляет собой совокупность математических объектов в виде структуры $\langle A, I, R, Pr \rangle$. Элементы данной структуры - множества объектов данной подсистемы: множество агентов-роботов A , информационное пространство I , множество ресурсов R , множество изделий (продуктов цифрового производства) Pr , агенты коммуницируют с помощью информационного пространства I .

Множество роботов-агентов (далее - агенты) $A = \{(a_1|q_1), (a_2|q_2), \dots, (a_n|q_1)\}$ реализует сборку изделия, q_i - значение, характеризующий уровень доступа агента к элементарным сообщениям информационного пространства, $0 \leq q_i \leq 1$. Каждый i -ый агент выполняет множество свойственных ему функций-команд $F_i = \{f_1, f_2, \dots, f_m\}$. Совокупность $\{F_i\}$ для каждого i -го агента формирует однозначно определенный алгоритм производства, в результате исполнения которого будет произведен определенный набор продукции $Pr = \{pr_1, pr_2, \dots, pr_v\}$. Детерминированность алгоритмов обуславливается тем, что с другой стороны произведенный продукт – результат некой функции $pr_i = f(A_i, F_i, R_i, I, t)$, косвенно зависящий от времени и набора элементарных сообщений информационного пространства I и напрямую зависящий от множества исходных ресурсов $R = \{r_1, r_2, \dots, r_s\}$. И если Q – пространство цифрового производства, тогда $f_{pr} : Q \rightarrow Pr$.

Информационное пространство, представленное на рис. 1, – это структура $\langle I, D \rangle$:

$I = \{i_1, i_2, \dots, i_l\}$ – множество элементарных информационных сообщений,

$D = \{d_1, d_2, \dots, d_l\}$, $0 \leq d_i \leq 1$ – множество элементов, характеризующих параметры доступа к i -му информационному сообщению.

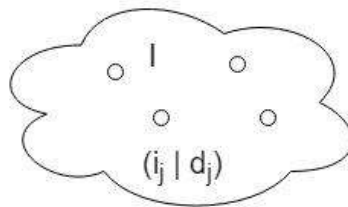


Рисунок 1. Информационное пространство

Информационное пространство может быть также представлено в виде отдельных кластеров информации, которые группируют информацию тематически:

$$I = I_{function} \cup I_{rules} \cup \dots \cup I_{client}$$

С другой стороны, информационное пространство представляется как совокупность подмножеств информационных сообщений сгруппированных положению в пространстве:

$$I = I_{main} \cup I_{interaction_{a,b}} \cup I_{agents}$$

Кластер I_{main} определяется как совокупность множеств информационных сообщений, сгруппированных по параметру доступа d , при том данный кластер – общий для всех агентов-роботов, агент a_i получает доступ к сообщению i_k , если выполняется условие $q_i \geq d_k$. Кластер $I_{interaction_{a,b}}$ определяет набор сообщений, передаваемых между агентами a и b . Последний кластер информации I_{agents} характеризует набор собственных сообщений k -го агента $\{i_k\}$.

Логистическая система представляется как множество автомобилей, выполняющих проезд из точки A (физического местоположения производства) до точки B (местоположения потребителя продукта).

Множество агентов-машин обозначим $Cars = \{car_1, car_2, \dots, car_n\}$. Автомобиль обладает набором параметров $p_i = \{status_i, location_i, resources_i\}$, где:

$$status_i = \begin{cases} 0, & \text{если } car_i \text{ свободен для выполнения задачи} \\ 1, & \text{если } car_i \text{ выполняет задачу} \end{cases},$$

$location_i = (x, y)$ – текущие координаты car_i , $resources_i$ – текущие ресурсы car_i .

Заключение

В ходе работы была разработана модель системы умного производства. Дальнейшая работа в области цифрового производства будет направлена на изучение свойств информации в информационном пространстве и влиянии на них деструктивного воздействия со стороны нарушителей.

ЛИТЕРАТУРА

1. Kang H.S. et al. Smart manufacturing: Past research, present findings, and future directions //International Journal of Precision Engineering and Manufacturing-Green Technology. – 2016. – Т. 3. – №. 1. – С. 111-128.
2. Lee J., Bagheri B., Kao H.A. A cyber-physical systems architecture for industry 4.0-based manufacturing systems //Manufacturing Letters. – 2015. – Т. 3. – С. 18-23.
3. Gerla M. et al. Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds //Internet of Things (WF-IoT), 2014 IEEE World Forum on. – IEEE, 2014. – С. 241-246.
4. Jedermann R. et al. Applying autonomous sensor systems in logistics—Combining sensor networks, RFIDs and software agents //Sensors and Actuators A: Physical. – 2006. – Т. 132. – №. 1. – С. 370-375.
5. Baloh M., Parent M. Modeling and model verification of an intelligent self-balancing two-wheeled vehicle for an autonomous urban transportation system //The conference on computational intelligence, robotics, and autonomous systems. – 2003. – С. 1-7.

Сведения об авторе

- | | |
|------------------------------|---|
| Мария Андреевна Усова | — студент; Университет ИТМО,
Факультет безопасности информационных технологий;
E-mail: gipurer@gmail.com |
| Илья Игоревич Вискнин | — научный сотрудник; Университет ИТМО,
Факультет безопасности информационных технологий;
E-mail: vixnin@cit.ifmo.ru |

Ссылка для цитирования: Усова М.А, Викснин И.И. Модель цифрового производства. // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 27—30.

SMART MANUFACTURE MODEL

M.A. Usova, I.I. Viksnin

ITMO University, 197101, St. Petersburg, Russia

The punctuation restoration problem arises in tasks of automatic correction of the scanned books and texts processing after the speech recognition system. In the article considered the existing approaches of the solution and their applicability in the context of rules of the Russian language. The choice of the best model and assessment of her accuracy when processing the unstructured text is made.

Keywords: punctuation restoration, neural network, sentence boundary disambiguation, natural language processing

Data on author

Maria A. Usova

— student; ITMO University,
Faculty of Information Security and Computer Technologies;
E-mail: gipurer@gmail.com

Ilya I. Viksnin

— researcher; ITMO University,
Faculty of Information Security and Computer Technologies;
E-mail: vixnin@cit.ifmo.ru

For citation: Usova M.A., Viksnin I.I. Smart manufacture model // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 27—30. (in Russian)

ТЕХНОЛОГИИ МУЛЬТИМЕДИА И КОМПЬЮТЕРНОЕ ЗРЕНИЕ

УДК 004.946

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ В ПСИХОКОРРЕКЦИИ

Д.А. Белова

Университет ИТМО, 197101, Санкт-Петербург, Россия

В статье анализируются возможности виртуальной реальности влиять непосредственно на психофизиологическое состояние человека. Сформулированы цель работы и задачи, которые должны быть выполнены для достижения цели. Указаны ожидаемые результаты планируемых исследований, а также области потенциального использования этих результатов.

Ключевые слова: виртуальная реальность, психотип, темперамент, психокоррекция, психокоррекционные методы

Введение

Технологии виртуальной реальности (VR) можно применять, чтобы воздействовать на психику человека с целью коррекции его состояния. VR позволяет программистам создавать реальность настолько правдоподобную реальность, что пользователь может полностью погрузиться в процесс исследования этого мира и выполнения запрограммированных в нем задач.

Таким образом, происходит отвлечение пользователя от окружающего мира и сильное воздействие на мозговую деятельность, что позволяет непосредственно влиять на его психофизиологическое состояние. Этим и объясняется актуальность данной работы. Поскольку виртуальная реальность обеспечивает достаточно высокий уровень погружения, такой метод можно считать более эффективным, чем прослушивание музыки или чтение книг для расслабления или выполнение упражнений для поднятия настроения.

Обзор предметной области

На сегодняшний день выделяют мобильные, десктопные и «независимые» системы виртуальной реальности, которые для работы с VR-приложениями предполагают использование или мобильных устройства, или персональных компьютеров, или же автономную работу соответственно [1].

Исследование сфер применения виртуальной реальности показало, что данные технологии можно эффективно использовать в сферах образования и здравоохранения, в том числе в психокоррекции: уже созданы приложения, которые используются при коррекции фобий, психических расстройств и состояния стресса [2, 3].

VR в психокоррекции используется при экспозиционной терапии (VRET – virtual reality exposure therapy). Например, офисы Медицинского центра виртуальной реальности в Калифорнии специализируются на лечении пациентов с тревожными расстройствами, обучении как военного, так и гражданского населения и расширении различных образовательных программ [4-6]. Еще одним наиболее распространенным направлением

являются ВР-приложения для релаксации. Например, было разработано приложение для HTC Vive, Samsung Gear VR и Oculus Rift – Guided Meditation VR – которое дает пользователю возможность выбрать подходящую ему успокаивающую виртуальную среду, в котором можно расслабиться и зарядиться энергией. Однако, все они не предполагают подхода, основанного на психотипе человека, а также еще не представлено приложения, которое было бы направлено на коррекцию монотонии.

Содержание исследования

В рамках проекта будет разработано приложение виртуальной реальности для психокоррекции на основе психотипа личности, которое будет использоваться для стабилизации состояния после монотонной или напряженной работы. Сценарий приложения будет разрабатываться под каждый психотип и психоэмоциональное состояние: монотония и напряженность.

Для того чтобы в дальнейшем сформировать эффективные сценарии стабилизации, было определено, что такое психотип и в чем особенность каждого, а также характеристики рассматриваемых состояний.

В рамках данного проекта психотип включает в себя типы высшей нервной деятельности (ВНД), предложенные И.П. Павловым, и типы личности, представленные Л.Н. Собчик. Таким образом, потенциальные пользователи могут быть разделены на координирующий, стимулирующий, содействующий и контролирующий типы. А суть монотонии и психоэмоциональной напряженности как функциональных состояний заключается в том, что в первом случае человеку необходимо взбодриться после однотипной работы, а во втором – успокоиться после понижения устойчивости психических и двигательных функций [7, 8].

Концепция проекта заключается в следующем:

1. Определяется психотип (темперамент) испытуемого;
2. Снимаются показания о нормальном состоянии человека и о состоянии человека после монотонной или напряженной работы до выполнения сценария приложения;
3. Применяется разработанная заранее сцена с алгоритмом действий под конкретный тип;
4. После работы с приложением снова снимаются показания;
5. Проводится сравнение результатов;
6. На выходе получается результат: прошел ли испытуемый стабилизацию успешно.

В данном проекте предполагается использование следующих методов исследования:

- сбор литературы и других источников информации по темам виртуальной реальности и психокоррекции;
- анализ предметной области;
- анализ методов психокоррекции;
- обобщение сформированной теоретической базы для выдвижения гипотез о психокоррекционных методах на основе психотипа человека и последующей практической реализации;
- сбор и анализ данных о состоянии испытуемых в покое и определение «идеальных» значений;
- проведение экспериментов: симуляция состояний монотонии и напряженности;
- тестирование испытуемых для определения психотипа (темперамента);

- сбор и анализ данных о состоянии испытуемых после монотонной/напряженной работы;
- проверка гипотез о воздействии ВР-сцен через создание сцен и апробации их испытуемыми с разными психотипами;
- тестирование приложения;
- сравнение полученных значений с ранее полученными “идеальными”;
- анализ полученных результатов.

Заключение

После завершения проекта появится готовая VR-система, которая может быть использована для проведения психокоррекционных процедур, подходящих для всех психотипов. Реализация проекта предполагает следующие результаты:

1. Подтверждение сформулированных гипотез о психокоррекционных методах, реализуемых технологиями виртуальной реальности;
2. Подтверждение стабилизации состояния субъекта после завершения сценария заявки.

Пользователями приложения могут стать все три группы потенциальных потребителей на рынке VR-технологий, выявленные Global Virtual Reality Association:

- обычные пользователи, которые используют VR-системы главным образом для развлечений;
- представители государственных органов, которые проявляют все больший интерес к VR-симуляторам так же, как к приложениям, способствующим привлечению внимания туристов к странам и их регионам;
- представители бизнеса, которые рассчитывают усовершенствовать производственный процесс с помощью данных технологий для предоставления нового вида услуг или средства связи с потребителями.

ЛИТЕРАТУРА

1. *LaValle S.M.* Virtual Reality. / S.M. LaValle – Cambridge University Press, 2017.
2. *Botella C., Quero S., Baños R.M., Perpiña C., Garcia-Palacios A., Riva G.* Virtual Reality and Psychotherapy. / C. Botella, S. Quero, R. M. Baños, C. Perpiña, A. Garcia-Palacios, G. Riva. // *Cybertherapy*, 99. – 2004.
3. *Brahnam S., Jain L. C.* Virtual Reality in Psychotherapy, Rehabilitation, and Neurological Assessment. / S. Brahnam, L. C. Jain. // *Advanced Computational Intelligence Paradigms in Healthcare*. – №6. – 2010.
4. *Eichenberg C.* Application of Virtual Realities in Psychotherapy: Possibilities, Limitations and Effectiveness. / C. Eichenberg. // In: Kim, J.-J, ed. – 2011.
5. *Riva G.* Virtual Reality in Psychotherapy: Review. / G. Riva. // *Cyberpsychology & Behavior*. – №3. – 2005.
6. *Riva G., Bacchetta M., Baruffi M., Molinari E.* Virtual Reality–Based Multidimensional Therapy for the Treatment of Body Image Disturbances in Obesity: A Controlled Study. / G. Riva, M. Bacchetta, M. Baruffi, E. Molinari. // *Cyberpsychology & Behavior* 4. – 2001.
7. *Собчик Л.Н.* Теория личности [Электронный ресурс] – URL: <http://psi-i.blogspot.com/1999/11/2003.html> (дата обращения: 10.10.2018).

8. Типы высшей нервной деятельности по И.П. Павлову [Электронный ресурс] – URL: <https://lektsii.org/15-46180.html> (дата обращения: 10.10.2018).

Сведения об авторе

Дарья Андреевна Белова

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: belovadaryaa@gmail.com

Ссылка для цитирования: *Белова Д.А.* Применение технологий виртуальной реальности в психокоррекции // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 31—34.

**USING VR TECHNOLOGIES AS A PART OF
PSYCHOCORRECTION PROCEDURES**

D.A. Belova

ITMO University, 197101, St. Petersburg, Russia

The paper analyzes the possibilities of VR to directly influence the psycho-physiological state of a person. The theses clearly formulate the purpose of the work and the tasks that must be accomplished to achieve the purpose. The expected results of the planned research are specified, as well as the areas of potential use of these results.

Keywords: virtual reality, psychotypes, temperament, psychocorrection, psychocorrection methods

Data on author

Daria A. Belova

— student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: belovadaryaa@gmail.com

For citation: *Belova D.A.* Using VR- technologies as a part of psychocorrection procedures // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems 2019. P. 31—34 (in Russian).

УДК 681.5.015.52

МЕТОДЫ ЭФФЕКТИВНОЙ ФИЛЬТРАЦИИ СИНТЕЗИРОВАННОГО ИЗОБРАЖЕНИЯ ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ РЕАЛИСТИЧНОГО РЕНДЕРИНГА РЕАЛЬНОГО ВРЕМЕНИ

А.Ф. Лемешев

Университет ИТМО, 197101, Санкт-Петербург, Россия

Исследуются существующие методы фильтрации изображений полученных реалистичным рендерингом в реальном времени. Фильтры реального времени должны учитывать это ограничение по времени, но при этом выдавать стабильный результат высокого качества. Рассмотренные методы соответствуют этим требованиям и по качеству изображения сравнимы с оффлайн-методами синтеза изображения. Для каждого алгоритма рассмотрены использованные техники, а также его слабые стороны. В заключении статьи выделены удачные решения алгоритмов фильтрации, а также дан общий вектор направления для эффективного решения задачи интерактивной фильтрации реалистичных изображений.

Ключевые слова: глобальное освещение, стохастическая трассировка лучей, физически корректный рендеринг, рендеринг в реальном времени, реконструкция изображения

Введение

Физически верное освещение сцены - это одна из основных проблем рендеринга в реальном времени. Для решения этой задачи обычно применяется стохастическая трассировка лучей. Этот подход основывается на случайном сэмплировании пикселя в каждом кадре. Основные сложности заключаются, прежде всего, в вычислительной сложности алгоритмов такой трассировки и алгоритмов фильтрации изображения на основе полученных данных. На текущий момент производительность трассировки лучей ограничена около 100 миллионов лучей в секунду на разрешении 1920x1080 при 30 кадрах в секунду, что дает нам по несколько лучей на пиксель. Производительность может быть еще ниже в динамических сценах. Это означает, что в обозримом будущем вычисление физического распространения света ограничено единственной трассировкой на пиксель. Фильтр реалистичного изображения обязан учитывать этот факт. Впрочем, как показали исследователи из Nvidia [1] этого более чем достаточно.

Обзор предметной области

На протяжении многих лет исследователи работали над физически верным освещением в реальном времени. Текущие разработки часто основаны на кэшировании предварительных вычислений трасс света [2] на поверхностях. Для того чтобы улучшить кэшированное освещение, используются методы, позволяющие включить дополнительные эффекты, такие как ambient occlusion, глянцевые отражения и мягкие тени [3]. Чаще всего данные решения строят приятное и правдоподобное изображение, но освещение получается не реалистичным.

Стохастические методы трассировки лучей позволяют правильно обработать некоторые сложные эффекты [4], такие как глубина резкости, размытие при движении, каустика, мягкие тени и глобальное освещение. Помимо большей реалистичности, при использовании трассировки лучей некоторые сложные эффекты проявляются сами собой, что может

снизить сложность итогового графического пайплайна. Это хорошо заметно в оффлайн рендеринге [5] и привело появлению многочисленных методов фильтрации, предназначенных для удаления остаточного стохастического шума. Цвикер [6] предоставил обзор данных методов.

Содержание исследования

В рамках исследования были проанализированы 3 современных алгоритма фильтрации физически верного освещения [1, 7, 8]. На основе этого анализа были выделены общие удачные решения и слабые стороны рассмотренных алгоритмов.

Удачные решения:

- Метод темпоральной аккумуляции позволяет сильно сократить шум за счет использования обработанных данных предыдущего кадра.
- Возможно сократить количество вычислений используя некоторые классические техники рендеринга, например, для вычисления теней использовать технику shadow mapping, а трассировку лучей использовать для глобального освещения
- Для того чтобы избежать артефактов при использовании темпоральной аккумуляции можно дополнительно обрабатывать те места для которых нет информации в предыдущем кадре (техника disocclusion detection).

Слабые стороны:

- Размытие и нечеткие границы - размытие происходит при большом количестве проходов пространственного фильтра, нечеткие границы можно убрать, используя граничную функцию на основе яркости
- Темпоральная аккумуляция обладает несколькими недостатками один из которых это эффект “призрачности” объекта (за объектом остается полупрозрачный след). От этого недостатка можно избавиться применив технику Gradient Estimation
- На данный момент алгоритмы фильтрации вычислительно сложны (не менее 5 миллисекунд на видеокarte NVIDIA TITAN X). Основные вычисления приходится на пространственные фильтры. Решить проблему с этим помогут все выше названное, а также использование разных фильтров для разных условий. Например, разные фильтры в зависимости от геометрии или глубины.

Дальнейшая разработка алгоритма фильтрации будет производиться с учетом этого анализа.

Заключение

Проанализированы современные алгоритмы фильтрации изображения и проанализированы подходы к оптимизации данных алгоритмов. На основе этого анализа были выделены удачные решения и общие недостатки. Принимая во внимание выделенные достоинства и недостатки, будет разработан метод фильтрации синтезированного изображения реалистичного рендеринга.

ЛИТЕРАТУРА

1. Schied C., Kaplanyan A., Wyman C., Patney A., Chaitanya C. R. A., Burgess J., Liu S., Dachsbacher C., Lefohn A., Salvi M. Spatiotemporal variance-guided filtering: real-time reconstruction for path-traced global illumination // Proceedings of High Performance Graphics, July 28-30, 2017, Los Angeles, California

2. Kajiya J.T. *The Rendering Equation*. // Computer Graphics (SIGGRAPH 86 Proceedings). 1986. Vol. 20. 143–150.
3. Ritschel T., Dachsbacher C., Grosch T., Kautz J. The State of the Art in Interactive Global Illumination. // Computer Graphics Forum 31, 1 (2012), 160–188.
4. Cook R. L., Porter T., Carpenter L. *Distributed Ray Tracing*. // Computer Graphics, Vol. 18. 1984. 137–145.
5. Keller A., Fascione L., Fajardo M., Georgiev I., Christensen P., Hanika J., Eisenacher C., Nichols G. *e Path Tracing Revolution in the Movie Industry*. // SIGGRAPH Courses. 2015.
6. Zwicker M., Jarosz W., Lehtinen J., Moon B., Ramamoorthi R., Rousselle F., Sen P., Soler C., Yoon S-E. Recent Advances in Adaptive Sampling and Reconstruction for Monte Carlo Rendering. // Computer Graphics Forum 34, 2 (2015), 667–681.
7. Dammertz H., Sewtz D., Hanika J., Lensch H. Edge-Avoiding A-Trous Wavelet Transform for fast Global Illumination Filtering. // In High Performance Graphics. 2010. 67–75.
8. Mara M., McGuire M., Bitterli B., Jarosz W. An efficient denoising algorithm for global illumination. // Proceedings of High Performance Graphics, July 28-30, 2017, Los Angeles, California

Сведения об авторе

Андрей Федорович Лемешев

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: andrewlemeshev@gmail.com

Ссылка для цитирования: Лемешев А.Ф. Методы эффективной фильтрации синтезированного изображения для решения проблемы реалистичного рендеринга реального времени // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 35—37.

METHODS OF EFFECTIVE FILTRATION OF A SYNTHESIZED IMAGE TO SOLVE THE PROBLEM OF REALISTIC REAL TIME RENDERING

A.F. Lemeshev

ITMO University, 197101, St. Petersburg, Russia

In this article, the author explored the existing methods of filtering the image obtained by realistic real-time rendering. Real-time filters should take into account the time limit of an interactive applications, but at the same time produce a stable, high-quality result. The methods, that mentioned in the article, fully correspond to these requirements and are comparable in quality with the offline image synthesis methods. For each algorithm, the used techniques and its weak points are considered. The conclusion of the article highlighted a successful solutions of filtering algorithms and mentioned a general direction to effectively solve the problem of interactive filtering of realistic images.

Keywords: global illumination, stochastic ray tracing, physically correct rendering, real-time rendering, image reconstruction

Data on author

Andrei F. Lemeshev

— student; ITMO University,
Faculty of Software Engineering and Computer Engineering;;
E-mail: andrewlemeshev@gmail.com

For citation: *Lemeshev A.F.* Methods of effective filtration of a synthesized image to solve the problem of realistic real time rendering // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 35—37 (in Russian).

УДК 004.4

ПОВЫШЕНИЕ КАЧЕСТВА ЗАПЕКАНИЯ ТЕКСТУРНЫХ КАРТ В ПРОЦЕССЕ СОЗДАНИЯ 3D КОНТЕНТА

А.В. Меженин, Ю.А. Трушина

Университет ИТМО, 197101, Санкт-Петербург, Россия

Рассматривается создание текстурных карт для низкополигональных 3D моделей и проблемы, возникающие в процессе их запекания. Запекание текстур – это процесс переноса детализации с высокополигональной 3D модели на ее низкополигональную копию с простой геометрией, который может применяться в игровой индустрии более эффективно. Обсуждаемые проблемы включают такие темы, как улучшение качества запекания с помощью групп сглаживания, настройка cage, разбиение полигональных групп, синхронизация касательной и др.

Ключевые слова: карта нормалей, нормали, запекание, текстуры, пайплайн, виртуальная и дополненная реальности

Введение

Причиной рассмотрения данной темы является важность темы создания 3D контента для современных игр [1]. Особенно применение технологии запекания карт нормалей актуально в связи с активным развитием технологий виртуальной и дополненной реальности, приложения для которых требуют большой оптимизации, и в то же время красивого качественного результата [2, 3]. Не менее важным оно является в прогрессивно развивающихся игровых движках, которые по своим возможностям приближаются к уровню качества классической покадровой визуализации. Тем не менее, до сих пор создатели 3D контента борются с возможными появлениями искажений перенесенной детализации картами нормалей, а также корректным отображением текстурных карт в используемых игровых движках.

Обзор предметной области

В области компьютерной 3D графики процесс создания (baking – «запекание») карты нормалей был и остается очень важным и активно используемым. Особенно применение технологии запекания карт нормалей актуально в связи с активным развитием технологий виртуальной и дополненной реальности, приложения для которых требуют большой оптимизации, и в то же время красивого качественного результата. Не менее важным оно является в прогрессивно развивающихся игровых движках, которые по своим возможностям приближаются к уровню качества классической покадровой визуализации.

Технология NormalMapping применяется в индустрии трехмерной графики для имитации детализации и неровностей поверхности трехмерных моделей. Она позволяет создавать иллюзию реалистичных поверхностей на низкополигональных моделях без увеличения количества их полигонов [4]. Карты нормалей представляют собой файлы изображений, каждый пиксель которых содержит информацию о направлениях поверхности высокополигональных 3D моделей. Таким образом, карта нормалей, накладываемая на низкополигональную модель, искусственно делает ее детализированной (рис. 1).

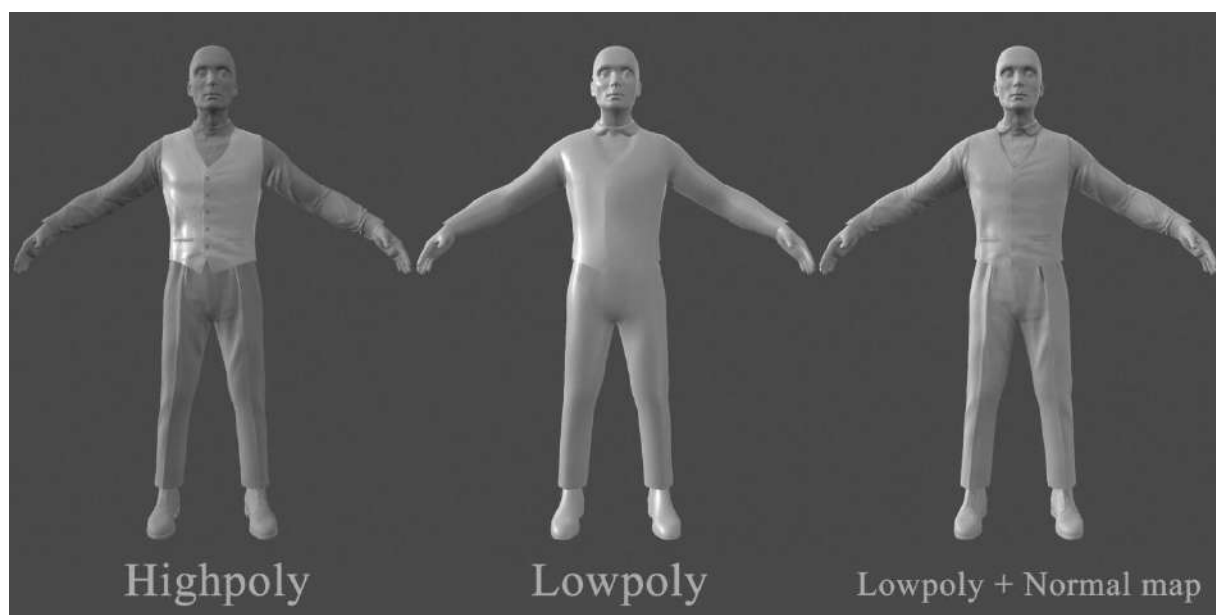


Рисунок 1. Высокополигональная модель, ее низкополигональная копия и низкополигональная модель с наложенной на нее картой нормалей

Содержание исследования

Создание карт осуществляется с помощью различных программных пакетов путем проецирования детализированной геометрии на низкодетализованную [5] (рис. 2).

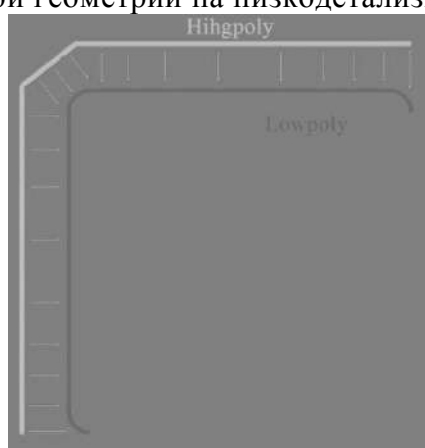


Рисунок 2. Проекционные лучи фиксируют изгиб геометрии на ребрах высокополигональной геометрии

Для результата проецирования очень важны разделения полигональной сетки моделей на группы сглаживания [6]. В разном программном обеспечении эту функцию осуществляют разные инструменты, например, в Autodesk Maya с помощью HardEdges производится расстановка ребер жесткости модели, а в Autodesk 3dsMax инструментом SmoothingGroups распределяются группы сглаживания для полигонов, границы которых являются жесткими.

Группа сглаживания обозначает усреднение нормалей вершин (vertexnormal – вектор, используемый для отображения поверхности в различных моделях освещения).

В этом случае программа пытается визуализировать куб, как шар: в пределах одной группы сглаживания направления нормалей всех вертексов усредняются, создавая, таким образом, плавный переход поверхности в пределах этой группы. Низкополигональная модель куба с одной группой сглаживания выглядит следующим образом (рис. 3).

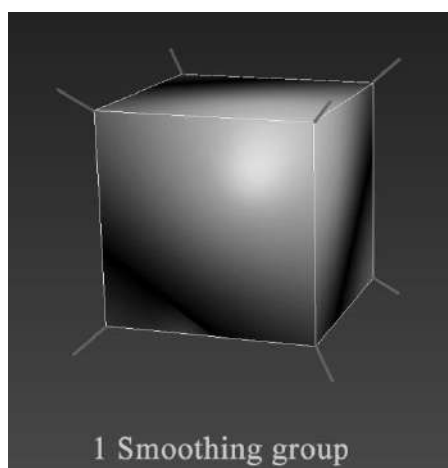


Рисунок 3. Куб с одной группой сглаживания и усредненными вертексными нормальными

В таком случае при попытке получить карту нормалей куба со сглаженными ребрами для достижения корректного результата (ровных плоскостей на гранях куба) карте нормалей придется компенсировать искажения сглаживания низкополигональной модели, что даст градиентную компенсацию на запеченном изображении нормалей (рис. 4).

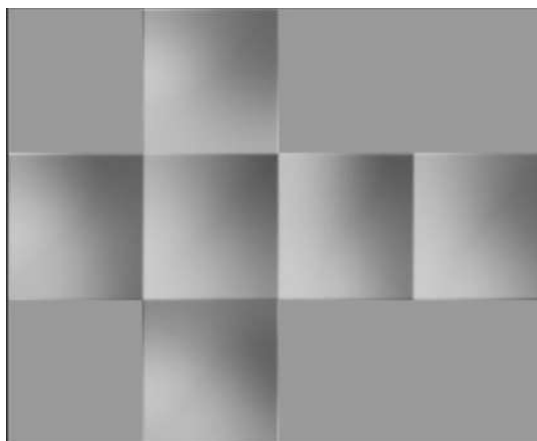


Рисунок 4. Карта нормалей куба с одной группой сглаживания, демонстрирующая компенсационные градиенты

Кроме того, что из-за сильных компенсаций в полученном результате все равно могут появляться артефакты (искажения). Пример: карта нормалей только частично компенсирует искажения низкополигонального куба с одной группой сглаживания, а на местах разреза UV-островов возникают артефакты в виде грубых швов на ребрах:

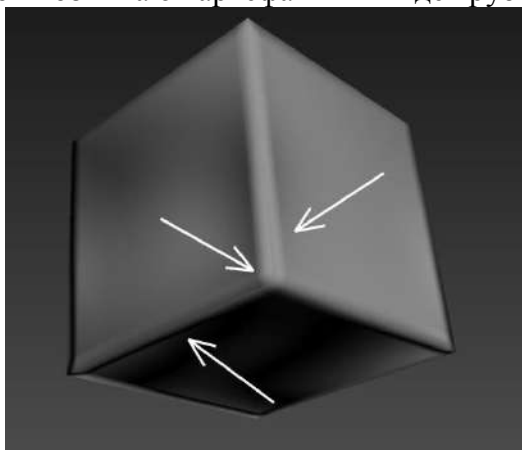


Рисунок 5. Искажения в виде градиентов на гранях куба и швов на границе UV-разрезов

Компенсационные градиенты на картах нормалей иногда затрудняют, а иногда делают невозможным их ручное редактирование при последующей работе. Это является большой проблемой, так как работа с картой нормалей не останавливается на процессе запекания – зачастую за ним следует ручная доработка в графических редакторах. Кроме того, на границах UVshells(частей UV-развертки полигональной сетки модели) при некорректном делении компенсационные градиенты могут порождать артефакты. А зачастую ограничение в допустимом в каждом конкретном случае разрешения изображения карт нормалей не всегда хватает, чтобы компенсировать сильное искажение (как правило, для запекания нормалей используется 8-битное изображение).

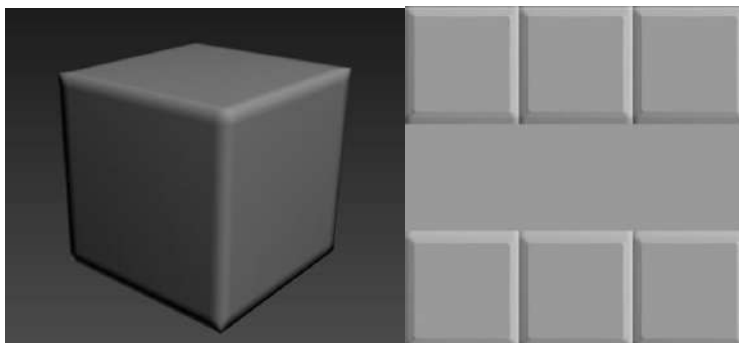


Рисунок 6. Куб с грамотно настроенными группами сглаживания и разверткой и его карта нормалей

Существуют различные способы избежать этих проблем. Самый популярный способ – это разделение полигональной сетки на группы сглаживания. Считается на практике, что между полигонами, находящимися под углом 90 градусов и меньше, обязательно следует назначать жесткие ребра, чтобы избежать видимых швов и артефактов, в остальных случаях – мягкие. Следовательно, распределять группы сглаживания необходимо в соответствии с углом геометрии.

Кроме того, на границах групп сглаживания низкополигональной модели, где, соответственно, назначены жесткие ребра, обязательно должны проходить разрезы в UV-развертке. Таким образом, при использовании групп сглаживания карты нормалей освобождаются от вынужденной компенсации, поэтому соответствующие градиенты ослабляются. А полученные карты нормалей, не загрязненные компенсационными градиентами, более пригодны и удобны для последующей ручной доработки [7].

Далее представлен результат грамотного распределения групп сглаживания и осуществления развертки для куба с аккуратно запеченными сглаженными ребрами без видимых искажений (рис. 6).

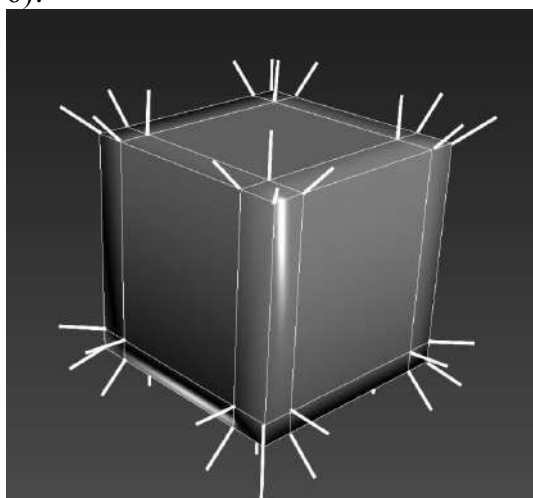


Рисунок 7. Добавление дополнительной геометрии, усредняющее вертексные нормали

Еще одним способом является добавление временной геометрии в той области модели, где при запекании карта нормалей вынуждена компенсировать неровности. В этом случае на низкополигональной модели добавляются дополнительные ребра для усреднения нормалей вертексов для запекания, после чего дополнительная геометрия удаляется, а выровненная с их помощью поверхность остается запеченной в карте нормалей (рис.7).

Ручное изменение направлений вертексов нормалей модели также является одним из способов устранения градиентных искажений, но он связан с рядом технических трудностей и ограничений и не всегда подходит.

Заключение

Рассмотренные методы позволяют разрешить ряд возникающих сложностей, связанных с градиентами на поверхности геометрии трехмерных моделей и компенсационных градиентов на запекаемых картах нормалей. В дальнейшем планируется продолжение рассмотрения методов устранения искажений перенесенной детализации на низкополигональных 3D моделях.

ЛИТЕРАТУРА

1. *Nguyen H.* (2007) GPU Gems 3, Addison-Wesley Professional
2. *Cohen J., Marc O., Manocha D.* (1998) Appearance-Preserving Simplification, Proceedings of the 25th annual conference on Computer graphics and interactive techniques, p. 115-122.
3. *Akenine-Moller T., Haines E., Homan N.* (2012), Real-time Rendering, Third Edition, A. K. Peters
4. *Абрамова О.Ф., Книжко А.В.* Исследование методов текстурирования ландшафта со сложным рельефом // NovaInfo.Ru (Электронный журнал). – 2016. – №55. [Электронный ресурс]: URL: <https://novainfo.ru/article/pdf/?nid=8950> (дата обращения: 10.01.2019).
5. *Bavoil L., Sainz M., Dimitrov R.* (2008) 'Image-space horizon-based ambient occlusion', Proceeding Siggraph'08 ACM SIGGRAPH 2008 talks.
6. *Меженин А.В.* Методы и средства распознавания образов и визуализации : Технология NormalMapping. – СПб: НИУ ИТМО, 2012. – С. 75-76.
7. Tips for Creating Perfect Normal Maps Every Time // Pluralsight. - 2014 - URL: <https://www.pluralsight.com/blog/tutorials/tips-creating-perfect-normal-maps-every-time>.

Сведения об авторе

- Александр Владимирович Меженин** — старший преподаватель; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: avmezhenin@corp.ifmo.ru
- Юлия Алексеевна Трушина** — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: yuliatr13@yandex.ru

Ссылка для цитирования: Меженин А.В., Трушина Ю.А. Повышение качества запекания текстурных карт в процессе создания 3D контента // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 38—43.

IMPROVING THE QUALITY OF BAKING TEXTURE MAPS IN THE PROCESS OF CREATION OF 3D CONTENT

A.V. Mezhenin, Y.A. Trushina

ITMO University, 197101, St. Petersburg, Russia

This thesis will look at texture-baking for games and the various problems that surrounds it. Texture baking is the process of transferring details from an often-times very dense highpolygonal details mesh to a lower detailed model that can be used in a game more efficiently. The problems discussed include topics such as failed tangent basis synchronization and improving bakes by using methods such as cages, smoothing groups splits, etc.

Keywords: Keywords: Normal map, normal, Texture-baking, , pipeline, virtual and augmented reality technologies

Data on author

- Alexandr V. Mezhenin** — senior lecturer; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: avmezhenin@corp.ifmo.ru
- Yulia A. Trushina** — student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: yuliatr13@yandex.ru

For citation: Mezhenin A.V., Trushina Y.A. Using holographic elements in the photorealistic rendering model // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 38—43 (in Russian).

УДК 004.428

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ ПРОГРАММНЫХ АЛГОРИТМОВ ДЛЯ ОПТИМИЗАЦИИ ИЗДЕЛИЙ В СРЕДЕ AUTODESK NETFABB

Д.А. Полякова, Ф.А. Перепелица

Университет ИТМО, 197101, Санкт-Петербург, Россия

Проанализированы методы решения проблемы адаптации трехмерных моделей для 3D-печати с использованием элементов генеративного проектирования и программных алгоритмов среды Autodesk Netfabb. Рассмотрены ограничения применения автоматических сценариев при оптимизации трехмерных моделей.

Ключевые слова: Autodesk Netfabb, подготовка 3D-моделей, оптимизация топологии, ремешивание, аддитивные технологии

Введение

На сегодняшний день Россия находится на начальном этапе совершенствования аддитивных технологий в сочетании с адаптированным к ним уровнем мышления пользователей [1]. Достичь высоких показателей в этих двух компонентах производства невозможно без применения комплексного подхода через осознание возможностей новых технологий и поисков соответствующих им подходов во всем — от дизайна 3D-модели до программного обеспечения компании. И это осознание становится стимулом не только к применению новых технологий, но и их совершенствованию.

Развитие концепции совершенствования технологий и методов аддитивного производства актуализировало проблему адаптации трёхмерных моделей для объемной печати.

Обзор предметной области

Во-первых, необходимо принимать во внимание ограничения аддитивного производства; в зависимости от используемых технологий и материалов изменяется процесс построения 3D-модели и её подготовки к выводу на печать. При использовании технологий аддитивного производства в большинстве случаев необходима генерация поддерживающих структур [2], не только для фиксации изделия и возможности осаждения материала на нависающих структурах, но и для компенсации термических влияний (теплоотвода). Однако, несмотря на ограничения аддитивного производства, 3D-печать позволяет создавать сложные бионические формы, которые невозможно получить с помощью традиционных субтрактивных методов [3].

Во-вторых, аддитивное производство позволяет воплотить в материале концепции генеративного дизайна, который стремительно набирает популярность в различных сферах производства: генеративный дизайн использует для синтеза геометрических форм математические алгоритмы и набор физических показателей, заданных пользователем компьютера.

Неадаптированные 3D-модели крайне негативно влияют на эргономику и производительность производства. Невостребованный испорченный пластик, который не будет в дальнейшем переработан и энергия, затраченная на работу принтера впустую,

заставляет задуматься о влиянии печатного производства на окружающую среду. Использование мелкодисперсных фракций металлических порошков, используемых при работе металлических принтеров для отработки технологии, негативно влияет на здоровье человека.

Цель исследования: провести анализ ограничений методов адаптации и оптимизации 3D-моделей для аддитивного производства с помощью внедрения в процесс подготовки 3D-моделей алгоритмов генеративного дизайна и программных алгоритмов среды Autodesk Netfabb.

Autodesk Netfabb включает несколько масштабных модулей, в которых включены программные алгоритмы подготовки 3D-моделей, сюда относятся: Support Module, Slice Commander, Part Repair, Lattices.

Проблематика исследования: несмотря на широкий функционал, модули не могут быть применены полностью в автоматическом режиме, требуется ручная доработка и вмешательство пользователя. Многие факторы оказывают негативное влияние на применение сценариев: это и исходные форматы файлов, тип построения файла (регулярная сетка или граничное представление), очередность выполнения операций, внедренная конфигурация оборудования и материальное решение. Генерация сложных параметрических структур (Mesh to BREP), создание решетчатых структур (Lattice Commander), заполнение файлов срезов штриховкой, создание траектории инструмента и откат перечисленных операций зачастую сопровождается потреблением значительного объема ресурсов памяти компьютера, возникновением коллизий или появлением нежелательных артефактов в геометрии модели.

Для решения описанных проблем выделены следующие задачи:

- Провести сравнительный анализ форматов импорта трехмерных файлов и соответствующих параметров модели.
- Аргументировать выбор оптимального формата и среды трехмерного моделирования для дальнейшей работы в программе Autodesk Netfabb.
- Проанализировать зависимость очередности выполнения операций по оптимизации моделей и качества полученной сетки.
- Провести сравнительный анализ топологий модели, полученной методом трехмерного сканирования до и после ремешивания.
- Определить факторы на уровне срезов трехмерной модели, подлежащие оптимизации.

Содержание исследования

Для выявления зависимости сеток трехмерных моделей от исходных форматов файлов был проведен эксперимент. В различных CAD-программах создавался эталонный образец «ГОСТ 11262-80. Пластмассы. Метод испытания на растяжение». Полученный образец сохранялся в одном из оригинальных форматов программы, а также экспортировался в формат .stl. Затем часть импортировалась в рабочую среду Autodesk Netfabb. Результаты анализа моделей приведены в Таблице 1.

Оптимальными форматами для оптимизации моделей в среде Autodesk Netfabb, являются: сетка в формате .stl, экспортированная из среды AutoCAD, а также твердотельная модель в формате .sldprt, созданная в программе SolidWorks. Иные форматы либо искажают оригинальные размеры модели при экспорте/конвертации файла, либо проводят редуцирование треугольников, тем самым снижая необходимую гладкость поверхностей оригинальной детали.

Очередность выполнения сценариев также влияет на итоговую триангуляцию сетки. Некоторые блоки программы автоматически предупреждают о порядке выполнения

операций. Так, модуль Support не может быть запущен при наличии ошибок в модели [4]. При генерации поддерживающих структур отсутствуют сценарии для их надежной репликации. После проведения вынужденных операций по редактированию части с поддержкой, Autodesk Netfabb предлагает повторную генерацию поддерживающих структур с учетом измененной геометрии. При изменении размеров модели может произойти пересечение поддерживающих структур с частью, что будет являться ошибкой. Поэтому генерацию поддержек рекомендуется делать на заключительном этапе позиционирования и масштабирования модели.

Таблица 1. Анализ моделей в среде Autodesk Netfabb

Формат файла	Triangles	Outbox (X*Y*Z)	Volume, cm ³	Area, cm ²	Type
SolidWorks					
.sldprt	1052	55*10*10	2.258	14.098	brep
.stl	2332	55*9.99*9.99	2.272	14.135	mesh
AutoCAD					
.fbx	648	55*10*9.90	2.250	14.071	mesh
.stl	900	55*10*10	2.258	14.098	mesh
Компас-3D					
.m3d	Формат не распознается программой				
.stl	404	55*10*10	2.246	14.083	mesh

При работе с разделами Hollowing и Latticing в модуле Lattice Assistant было выявлено, что наилучшие результаты достигаются, когда оригинальная сетка состоит из треугольников правильной формы, например, созданных путем Quad Remeshing [5]. Автоматический ремонт моделей не всегда дает удовлетворительный результат даже при применении расширенных сценариев ремонта. Тривиальное закрытие отверстий приводит к неестественной плоской «заплатке», которая в ряде случаев выглядит неестественно. Обязательна ручная доработка.

Для исследования эффективности алгоритмов ремешивания использовалась обработка различных моделей. Выявлено, что ремешивание параметрических сеток неэффективно, поскольку изделие сильно деформируется по форме и возникает множество артефактов. Наилучшие результаты были получены при ремешивании моделей, полученных методом трехмерного сканирования. В эксперименте использовалась трехмерная модель *Carcinoplax_Suruguensis.stl*. До ремешивания модель содержала 1801090 треугольников, и множественные артефакты на поверхности сетки (Рис. 1), после ремешивания модель содержит гладкую поверхность и оптимизированное число треугольников (74082).

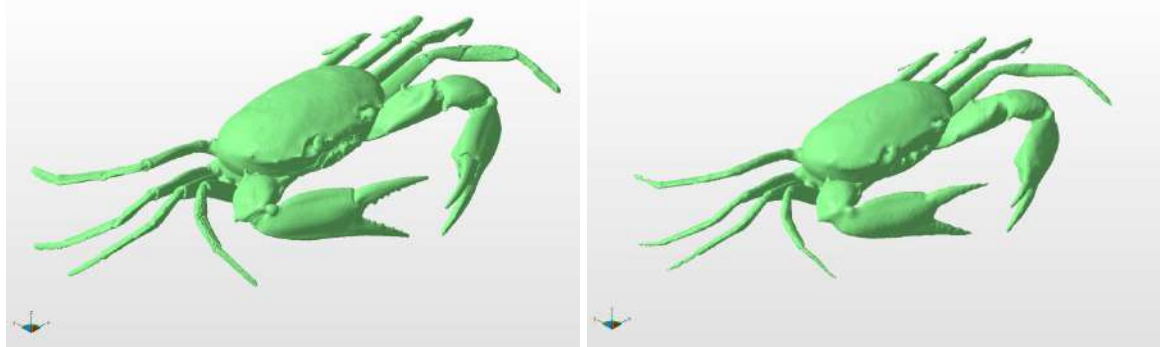


Рисунок 1. *Carcinoplax_Suruguensis* до ремешивания (слева), после ремешивания (справа)

Избыточная триангуляция сетки генерирует высокое количество опорных точек при слайсинге модели, что оказывает негативное влияние на генерацию дальнейшей траектории лазера и расходу филамента. Функция оптимизации количества точек (Reduce point) позволяет получить оптимизированный срез аналогичной формы (Рис. 2). Неверное

значение параметра максимальной деформации, приведет к потере плавных форм и искажению геометрии.

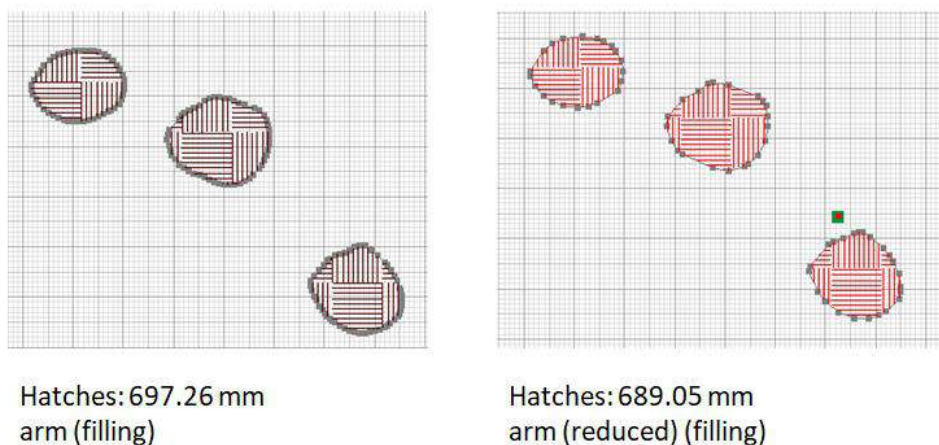


Рисунок 2. Генерация заполнения для неоптимизированного контура (слева), после проведения редуцирования (справа)

От неоптимизированного количества опорных точек напрямую зависит генерация внутреннего заполнения среза. Избыточность среза приводит к сложной генерации траектории лазера и повышенному расходу филамента на всей высоте стека.

Заключение

Выявлена программная зависимость триангулярных сеток от программной среды CAD-моделирования. Определено различие триангуляции .stl-форматов, экспортированных из различных программ 3D-моделирования. Оптимальными форматами для обработки файлов в программе Autodesk Netfabb являются форматы .stl (экспорт файла из программы AutoCAD), .sldprt (сохранение в оригинальном формате программы SolidWorks).

На качество исходной триангулярной сетки зачастую оказывает влияние очередность выполнения сценариев. Так, создание решетчатых структур, полостей и перфораций занимает меньше времени при предварительном ремешивании сетки путем Quad Remeshing. Применение автоматических сценариев сопровождается необходимостью ручной доработки.

Алгоритмы ремешивания не рекомендуется применять к параметрическим деталям и деталям серийного производства. Алгоритмы ремешивания не отслеживают исходную геометрию детали. Наилучшие результаты ремешивания наблюдаются при использовании сценариев на моделях, полученных методом трехмерного сканирования.

Высокая триангуляция сетки сопровождается избыточной генерацией опорных точек на внешнем контуре среза, что в дальнейшем окажет негативное влияние на создание внутреннего заполнения. Грамотное применение операции редуцирования точек позволит оптимизировать контур среза.

ЛИТЕРАТУРА

1. Сухочев Г.А., Технология машиностроения. Аддитивные технологии в подготовке производства наукоемких изделий / Г.А. Сухочев, С.Н. Коденцев, Е.Г. Смольяникова – Воронеж: Воронежский гос. технический ун-т, 2013. – 222 с.
2. Perry Cain, Supports in 3D Printing: A technology overview // 3D Hubs, 2017.
3. Краснов А.А., Смоленцев Е.В. Аддитивное и субтрактивное производство // Инновационная наука. 2016. № 12-2.

4. Netfabb Professional 5.2 User Manual // Netfabb GmbH, July 29, 2014.
5. Autodesk Netfabb User Manual For version 2017 // Netfabb GmbH, November 2016.

Сведения об авторе

Дарья Александровна Полякова	—	студент; Университет ИТМО, Факультет программной инженерии и компьютерной техники; E-mail: pooda6a@gmail.com
Филипп Александрович Перепелица	—	старший преподаватель; Университет ИТМО, Факультет программной инженерии и компьютерной техники; E-mail: perepelitcafa@niuitmo.ru

Ссылка для цитирования: Полякова Д.А., Перепелица Ф.А. Особенности использования программных алгоритмов для оптимизации изделий в среде Autodesk Netfabb // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 44—48.

**FEATURES OF THE USE OF SOFTWARE ALGORITHMS AND
TECHNOLOGIES OF GENERATIVE DESIGN
TO OPTIMIZE PRODUCTS IN AUTODESK NETFABB**

D.A. Polyakova, P.A. Perepelitca

ITMO University, 197101, St. Petersburg, Russia

Analyzed methods for solving the problem of adapting three-dimensional models for 3D printing using elements of generative design and software algorithms of the Autodesk Netfabb environment. The limitations of the use of automatic scripts in optimizing three-dimensional models are considered.

Keywords: Autodesk Netfabb, preparation of 3D-models, topology optimization, remesh, additive technologies

Data on author

Darya A. Polyakova	—	student; ITMO University, Faculty of Software Engineering and Computer Engineering; E-mail: pooda6a@gmail.com
Philipp A. Perepelitca	—	senior Lecturer; ITMO University, Faculty of Software Engineering and Computer Engineering; E-mail: perepelitcafa@niuitmo.ru

For citation: Polyakova D.A., Perepelitca P.A. Features of the use of software algorithms and technologies of generative design to optimize products in Autodesk Netfabb // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 44—48 (in Russian).

УДК 004.93

ПРИМЕНЕНИЕ ИНС ТИПА «ПЕРСЕПТРОН» В МОРФОЛОГИЧЕСКОЙ ДИАГНОСТИКЕ ПАПИЛЛЯРНОГО РАКА ЩИТОВИДНОЙ ЖЕЛЕЗЫ

**А.А. Смирнов¹, Е.М. Трунин¹, И.В. Исаев², И.В. Чижиков², А.Л. Овсепьян³,
О.Б. Бегишев¹, Л.Л. Мурт¹, Ш.Ш. Кудлахмедов¹, А.И. Назмиев¹, П.А. Сизов¹,
Д.Б. Бабаев¹**

1 - Северо-Западный государственный медицинский университет им. И.И. Мечникова, Санкт-Петербург, Россия.

2 - Университет ИТМО, Санкт-Петербург, Россия.

*3 - Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия.*

В статье описаны этапы разработки, обучения и апробации нелинейного программного обеспечения на основе искусственной нейронной сети типа «персептрон» для диагностики папиллярного рака щитовидной железы. Подробно описан оригинальный алгоритм перевода гистологического изображения в абстрактную форму планарного графа и последующий его анализ с применением разработанного программного обеспечения.

Ключевые слова: щитовидная железа, папиллярный рак, искусственная нейронная сеть

Введение

Хирургическая наука в настоящий момент проходит период бурного развития, следуя параллельным курсом с научно-техническим прогрессом и активно перенимая его передовые достижения. Одним из таких передовых достижений является применение систем технического зрения, среди которых важное место в медицинских исследованиях относительно недавно заняли различные типы искусственных нейронных сетей, которые представляют собой кибернетические аналоги нейронных сетей живых организмов и могут решать, наравне с последними, сложнейшие задачи, связанные с обработкой и анализом огромного объема информации.

Обзор предметной области

Папиллярный рак (ПР) занимает первое место среди всех злокачественных новообразований щитовидной железы, составляя приблизительно 78-85% всех случаев этого заболевания. Поскольку эта форма рака щитовидной железы растет крайне медленно и метастазирует преимущественно лимфогенным путём в регионарные лимфатические узлы, необходимость своевременной диагностики ПР не вызывает сомнений. В настоящее время, для уточнения диагноза возможно выполнять забор фрагмента ткани узла, подозрительного на ПР, с помощью устройства для взятия пробы ткани щитовидной железы [1] с последующим его гистологическим исследованием. С этой целью нами разработан способ морфологической диагностики ПР щитовидной железы с применением программного обеспечения (ПО) на основе искусственной нейронной сети (ИНС) типа «персептрон» [4,6].

Целью исследования является разработка и апробация программного обеспечения на основе ИНС типа «персептрон» для морфологической диагностики папиллярного рака щитовидной железы.

Задачи исследования:

1. Произвести отбор и фотофиксацию гистологических препаратов нормального гистологического строения ткани щитовидной железы и папиллярного рака щитовидной железы.
2. Разработать программное обеспечение на основе ИНС типа «персептрон» с оригинальным алгоритмом преобразования гистологического изображения в планарный граф.
3. Выполнить обучение ИНС, используя маркированные учебные примеры, полученные в процессе фотофиксации гистологических препаратов и дальнейшего преобразования изображений.
4. Экспериментально проверить и установить эффективность применения программного обеспечения на основе искусственной нейронной сети типа «персептрон».

Содержание исследования

На первом этапе работы были отобраны две группы гистологических препаратов (рис. 1): первая - с нормальным строением ткани щитовидной железы ($n=320$) и вторая - с ПР щитовидной железы ($n=283$). После этого, на увеличении $\times 20$, была произведена фотофиксация с помощью микроскопической цифровой камеры LEVENHUK C310 2-5 различных непересекающихся локусов каждого из стеклопрепаратов. Фотофиксация локусов нормального строения щитовидной железы производилась при исследовании стеклопрепаратов с такими заключениями как: нетоксический многоузловой зоб, аутоиммунный тиреоидит, диффузно-узловой токсический зоб и т.д. (табл. 1). Каждое из полученных изображений, после описанной ниже программной обработки, представляло собой маркированный учебный пример, который использовался в последующем обучении нейросети.

ИНС, которую мы использовали для обработки изображений, представляет собой персептрон прямого распространения с 2-мя скрытыми слоями (рис. 2). Данная конфигурация ИНС хорошо подходит для выполнения задачи классификации образов, при которой необходимо соотнести входной сигнал, представляющий собой физический объект с некоторой предопределённой категорией или классом. Обучение данной нейросети происходит методом обратного распространения ошибки.

Таблица 1. Распределение исследуемых стеклопрепаратов по видам патологии щитовидной железы и общему количеству исследуемых локусов

Диагноз	Количество стеклопрепаратов	Кол-во локусов
Нетоксический многоузловой зоб.	90	294
Узловой нетоксический зоб. Компрессионный синдром.	42	126
Узловой нетоксический зоб.	117	393
Аутоиммунный тиреоидит, узловая форма.	32	95
Диффузно-узловой токсический зоб.	14	37
Многоузловой токсический зоб.	9	28
Диффузный токсический зоб.	16	35
Папиллярная карцинома ЩЖ.	283	761
ИТОГО	603	1769

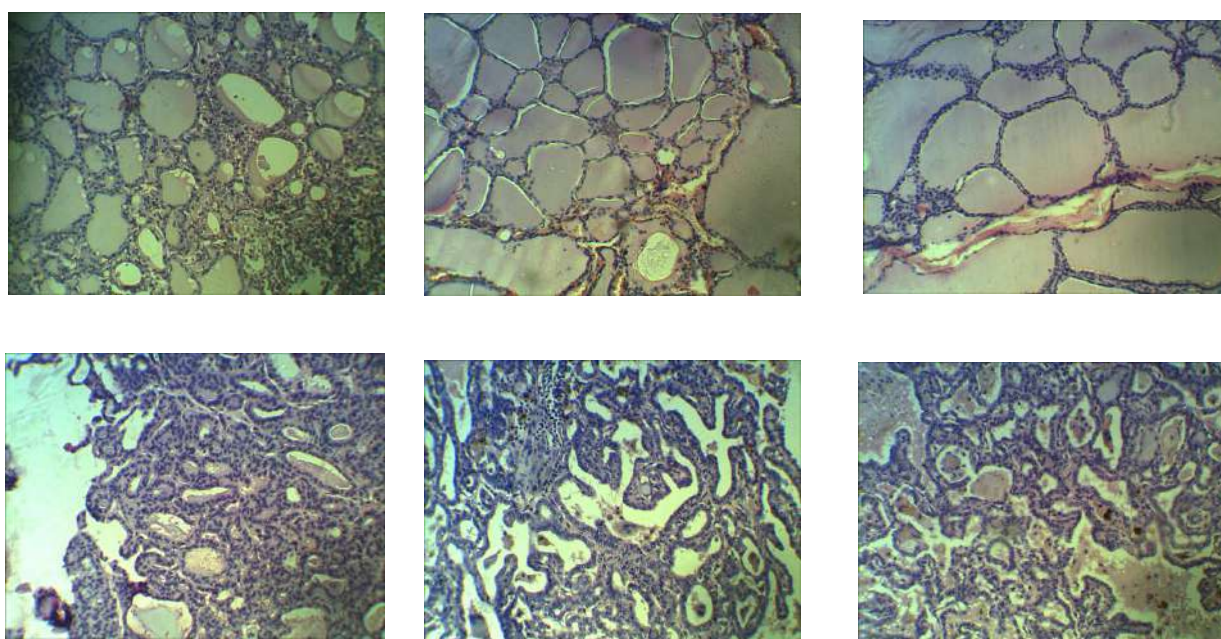


Рисунок 1. Фотографии гистологических препаратов щитовидной железы (верхний ряд – нормальное гистологическое строение щитовидной железы, нижний ряд – папиллярный рак)

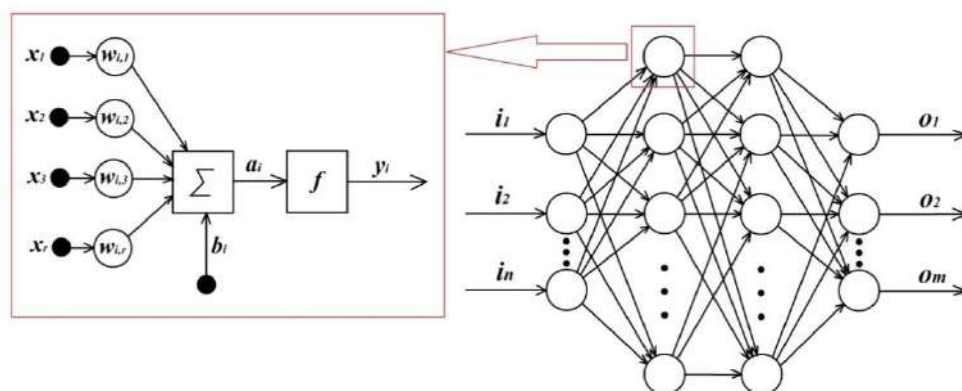


Рисунок 2. Принципиальная схема устройства ИНС типа «персептрон» с двумя скрытыми слоями: слева – структура искусственного нейрона, состоящего из входных сигналов (x), синаптических весов (w), сумматора (Σ), порога (b), активационной функции (f) и выходного сигнала (y); справа – структура нейросети,

построенной из отдельных искусственных нейронов и имеющей входной слой (i), два скрытых слоя и выходной слой (o)

При выполнении этого исследования, мы выбрали один из самых популярных методов «обучения» ИНС, так называемое «обучение с учителем» (supervised learning), которое подразумевает изменение синаптических весов на основе набора маркированных учебных примеров. Каждый такой пример представляет собой входной сигнал и соответствующий ему желаемый отклик. При каждом предъявлении маркированного учебного примера, ИНС модифицирует синаптические веса для минимизации расхождений желаемого выходного сигнала, формируемого сетью согласно выбранному статистическому критерию. Используя предварительно отобранные снимки, после перевода их по указанному ниже алгоритму в набор плоских графов, были произведены обучение и корректировка ИНС.

Корректировка параметров сети производилась с учетом обучающего вектора и сигнала ошибки, т.е. разницы между желаемым откликом и текущим откликом нейронной сети. Корректировка синаптических весов выполняется с целью имитации нейронной сетью «поведения» учителя и минимизации расхождений желаемого выходного сигнала [4]. Оценка погрешности работы ИНС была рассчитана методом наименьших квадратов (МНК, англ. Ordinary Least Squares, OLS).

Второй этап работы заключался в преобразовании полученных гистологических изображений в абстрактную математическую форму планарных графов с помощью разработанных нами программных модулей THYROCULUS 1.0 и 2.0 [2, 3], работающих на основе оригинального алгоритма. Предварительная обработка изображения включала в себя изменение настроек алгоритма определения ядер с применением цветовой модели HSV. В ходе этого этапа ПО выделяет на изображении гистологического препарата все области, имеющие оттенки фиолетового цвета в заданном диапазоне. В результате, происходило преобразование изображения таким образом, чтобы на нем оставались только лишь ядра клеток. Далее, каждую группу пикселей (ядра) обозначали как отдельный объект. После этого ПО определяло геометрические центры всех объектов и отображало их на белом фоне с сохранением исходных пространственных взаимоотношений, отмечая его точкой с известными координатами. Таким образом, гистологическое изображение преобразуется в набор точек на плоскости. После этого, ПО определяло точки, ближайшие к каждой точке на плоскости с соединяло их прямыми линиями, завершая, таким образом, формирование планарного графа (рис. 3). Таким образом, в виде плоского графа, моделируется цитоархитектоника в изучаемом плоском срезе. Важнейшим качественным переходом на этом этапе является то, что образ на изображении гистологического препарата преобразуется в плоский (планарный) граф – абстрактный математический объект, представляющий собой множество вершин и соединяющих их рёбер. Кроме того, на этом этапе производилась настройка и корректировка алгоритма перевода гистологических изображений в планарные графы.

Кроме большого теоретического значения, перевод гистологического изображение в форму планарного графа имеет также важное значение практическое, как способ унификации гистологических препаратов, изготовленных в разных лабораториях, и имеющих по этой причине незначительные отличия. Кроме того, переход к форме планарного графа значительно уменьшает «шум» изображения, полностью нивелируя незначительные в диагностическом плане объекты.

Третий этап – это непосредственное обучение ИНС, заключавшееся в предъявлении ей полученных планарных графов. С целью проверки эффективности выполненного обучения, ИНС предоставлялись ранее незадействованные в её обучении снимки – по 50 из каждой группы («нормы» и патологии).

При предъявлении контрольных (ранее незадействованных в «обучении») снимков гистологических препаратов с нормальным строением ткани щитовидной железы ($n=50$) и с ПР щитовидной железы ($n=50$) (рис. 4), разработанное ПО на основе ИНС типа «персептрон» дало 92% правильных ответов по папиллярному раку щитовидной железы и 84% совпадений по образцам нормальной ткани щитовидной железы. Суммарная точность работы ИНС составила, таким образом, 88%.

Заключение

Разработанные в ходе данной исследовательской работы программные модули на основе ИНС типа «персептрон» способны эффективно распознавать и различать нормальное гистологическое строение ткани щитовидной железы и папиллярного рака щитовидной железы. Таким образом, разработанное ПО может быть использовано в работе патологоанатомических отделений городских больниц в качестве эффективной системы поддержки принятия решений (СППР). С целью усовершенствования разработанного нами СППР и увеличения его эффективности, мы планируем значительно расширить диагностические возможности разработанного ПО путём увеличения количества категорий распознавания для ИНС, а также изменения формы и характера вводимых в ИНС данных.

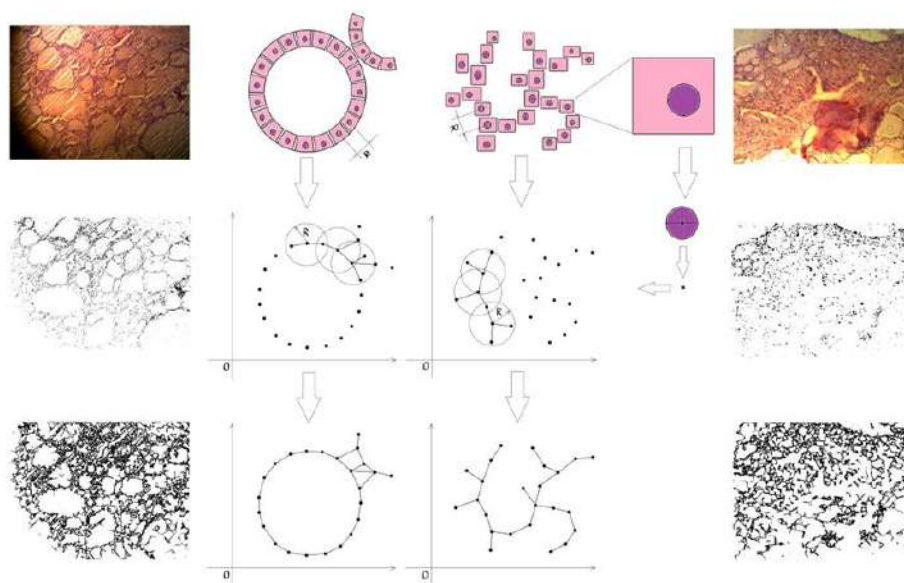
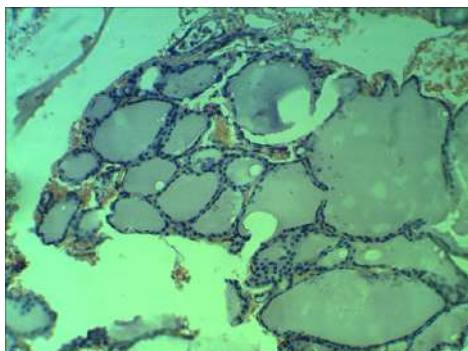
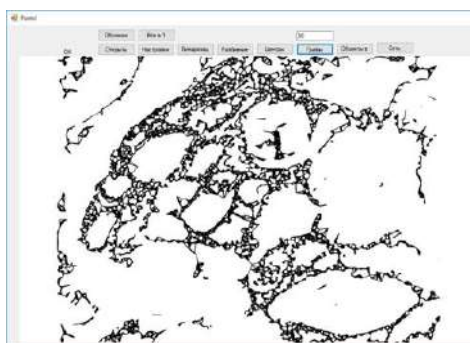
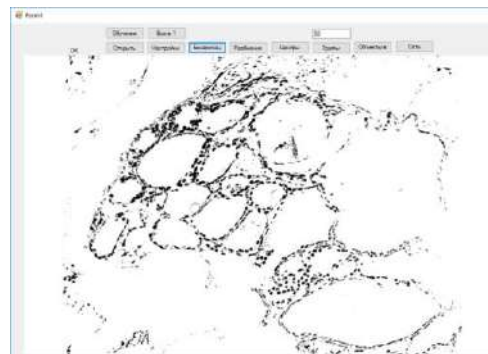


Рисунок 3. Этапы формирования планарного графа на основе фотографии гистологического среза (в центре – упрощенная и обобщенная схема, слева и справа – этапы преобразования фотографий гистологических срезов в планарные графы)

1



2



3

4

Рисунок 4. Этапы работы разработанного программного обеспечения на примере образца нормального гистологического строения щитовидной железы: 1 – фотография гистологического изображения нормального гистологического строения щитовидной железы, 2 – бинаризация (этап поиска ядер и выделение их в отдельный объект), 3 – сформированный планарный граф, 4 – выходные данные ПО (ответ) [2,3]

ЛИТЕРАТУРА

1. Пат. 2551957 Российская Федерация МПК, А61В 10/00 Устройство для взятия пробы ткани щитовидной железы [текст] / Трунин Е.М., Татаркин В.В., Смирнов А.А., Шульга В.П.; - № 20141180334/14; заявл. 05.05.2014; опубл. 10.06.2015, Бюл. № 16. – 14 с.: ил.
2. Смирнов А.А., Чижиков И.В., Исаев И.В. и др. Искусственная нейронная сеть типа «персептрон» для морфологической диагностики папиллярного рака щитовидной железы THYROCULUS 1.0 / Свидетельство о государственной регистрации ПрЭВМ, рег. №2018662177 от 27.09.2018. – М.: Роспатент, 2018.
3. Смирнов А.А., Чижиков И.В., Исаев И.В. и др. Искусственная нейронная сеть типа «персептрон» для морфологической диагностики папиллярного рака щитовидной железы THYROCULUS 2.0 / Свидетельство о государственной регистрации ПрЭВМ, рег. №2018662178 от 27.09.2018. – М.: Роспатент, 2018.
4. Хайкин, С. Нейронные сети: полный курс / Саймон Хайкин; [пер. с англ. д.т.н. Н.Н. Куссуль, к.т.н. А.Ю. Шелестова]. - 2-е изд. - Москва [и др.]: Вильямс, 2006. - 1103 с.
5. Lorensen W. E., Cline H. E. Marching cubes: A high resolution 3D surface construction algorithm //ACM siggraph computer graphics. – ACM, 1987. – Т. 21. – №. 4. – С. 163-169.

6. J. Jiang, P. Trundle, J. Ren, Medical image analysis with artificial neural networks, Computerized Medical Imaging and Graphics, Volume 34, Issue 8, 2010, Pages 617-631, ISSN 0895-6111, <https://doi.org/10.1016/j.compmedimag.2010.07.003>.

Сведения об авторе

Александр Александрович Смирнов	—	доцент; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра оперативной и клинической хирургии с топографической анатомией; E-mail: savmeda@yandex.ru
Евгений Михайлович Трунин	—	заведующий кафедрой; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра оперативной и клинической хирургии с топографической анатомией; E-mail: etrunin@mail.ru
Олег Бурханович Бегишев	—	доцент; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра оперативной и клинической хирургии с топографической анатомией
Шакир Шавкатович Кудлахмедов	—	врач-ординатор; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра офтальмологии; E-mail: shakir47@yandex.ru
Лариса Леонидовна Мурт	—	заведующая отделением, НУЗ "Дорожная клиническая больница" ОАО "РЖД", отделение морфологической диагностики
Артур Левонович Овсепьян	—	студент, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Кафедра биотехнических систем; E-mail: aqturovsepyan@gmail.com
Павел Алексеевич Сизов	—	студент; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра оперативной и клинической хирургии с топографической анатомией; E-mail: pavelsizov@mail.ru
Азат Ильдусович Назмиев	—	студент; Северо-Западный государственный медицинский университет им. И.И. Мечникова, Кафедра оперативной и клинической хирургии с топографической анатомией; E-mail: azat.nazmiev1995@gmail.com
Илья Владимирович Исаев	—	ассистент, Университет ИТМО, Факультет программной инженерии и компьютерной техники; E-mail: ivisaev@corp.ifmo.ru

Ссылка для цитирования: Смирнов А.А., Трунин Е.М., Исаев И.В., Чижиков И.В., Овсепьян А.Л., Бегишев О.Б., Мурт Л.Л., Кудлахмедов Ш.Ш., Назмиев А.И., Сизов П.А., Бабаев Д.Б. Применение инс типа «персептрон» в морфологической диагностике папиллярного рака щитовидной железы // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 49—56.

THE USE OF ANN OF THE «PERCEPTON» TYPE IN THE DIAGNOSIS OF PAPILLARY THYROID CARCINOMA

A.A. Smirnov¹, E.M. Trunin¹, I.V. Isaev², I.V. Chizhikov², A.L. Ovsepyan³, O.B. Begishev¹, L.L. Murt¹,

S.S. Kudlakhmedov¹, A.I. Nazmiev¹, P.A. Sizov¹, D.B. Babaev¹

1 - North-Western State Medical University named after I.I. Mechnikov, 191015, St. Petersburg, Russia.

2- ITMO University, 197101, St. Petersburg, Russia

3 - Saint Petersburg Electrotechnical University "LETI", 197022, St. Petersburg, Russia.

The article describes the stages of development, training and testing of a nonlinear software based on an artificial neural network of the "perceptron" type for the diagnosis of papillary thyroid cancer. The original algorithm for converting a histological image into an abstract form of a planar graph and its subsequent analysis using the developed software are described in detail.

Keywords: thyroid gland, papillary thyroid carcinoma, artificial neural network

Data on author

Alexandr A. Smirnov	—	docent; North-Western State Medical University named after I.I. Mechnikov, Department of Operative and Clinical Surgery with Topographical Anatomy; E-mail: savmeda@yandex.ru
Eugeny M. Trunin	—	head of the department; North-Western State Medical University named after I.I. Mechnikov, Department of Operative and Clinical Surgery with Topographical Anatomy; E-mail: etrunin@mail.ru
Oleg B. Begishev	—	docent; North-Western State Medical University named after I.I. Mechnikov, Department of Operative and Clinical Surgery with Topographical Anatomy
Shakir S. Kudlakhmedov	—	resident doctor; North-Western State Medical University named after I.I. Mechnikov, Department of Ophtalmology; E-mail: shakir47@yandex.ru
Larisa L. Murt	—	head of the department, Railway clinical hospital, Department of the morfological diagnosis
Artur L. Ovsepyan	—	student, Saint Petersburg Electrotechnical University "LETI", Department of Bioengineering Systems; E-mail: aqturovsepyan@gmail.com
Pavel A. Sizov	—	student; North-Western State Medical University named after I.I. Mechnikov, Department of Operative and Clinical Surgery with Topographical Anatomy; E-mail: pavelsizov@mail.ru
Azat I. Nazmiev	—	student; North-Western State Medical University named after I.I. Mechnikov, Department of Operative and Clinical Surgery with Topographical Anatomy; E-mail: azat.nazmiev1995@gmail.com
Ilya V. Isaev	—	assistant, ITMO University, Faculty of Software Engineering and Computer Systems; E-mail: ivisaev@corp.ifmo.ru

For citation: Smirnov A. A., Trunin E.M., Isaev I.V., Chizhikov I.V., Ovsepyan A.L., Begishev O.B., Murt L.L., Kudlakhmedov S.S., Nazmiev A.I., Sizov P.A., Babaev D.B. The use of ANN of the «perceptron» type in the diagnosis of papillary thyroid carcinoma // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 49—56(in Russian).

УДК 004.922

ПОДХОДЫ К АВТОМАТИЗАЦИИ ОБРАБОТКИ ИЗОБРАЖЕНИЯ ФОРМАТА 360 ГРАДУСОВ

И.А. Стоколяс

Университет ИТМО, 197101, Санкт-Петербург, Россия

Развитие технологий производства фото- и видеоаппаратуры для съемки широкоформатных изображений выявило потребность в автоматизации обработки сферических панорам. В процессе исследования изучены основные существующие программные средства и подходы к автоматизации обработки сферических панорам, выявлены их преимущества и недостатки. Было разработано программное средство, учитывающее потребности пользователя при работе с программой. Результаты экспериментов свидетельствуют о том, что разработанная программа успешно автоматизирует работу пользователя.

Ключевые слова: BoofCV, OpenCV сферическая панорама, эквидистантная панорама, обработка изображений, калибровка камеры, графический пользовательский интерфейс, автоматизация

Введение

Основными причинами медленной работы пользователя с различными программами обработки изображений являются сложные и перегруженные графические интерфейсы программ, а также используемые алгоритмы, которые требуют большого количества настроек со стороны пользователя. Эти причины вызывают потребность в автоматизации действий пользователя.

Целью работы является автоматизация действий пользователя по обработке исходных изображений с камер формата 360 градусов. Для достижения цели были поставлены следующие задачи:

- анализ существующих программных средств и подходов к автоматизации создания сферических панорам из фотографий со специализированных камер;
- разработка программы с учетом данных, полученных в ходе анализа;
- проведение эксперимента для оценки качества автоматизации.

Обзор предметной области

Специализированные камеры, позволяющие вести съемку формата 360 градусов, имеют два объектива формата «Рыбий глаз», что позволяет с помощью одного снимка получить достаточно данных для создания сферической панорамы. Таким образом пользователь получает исходное изображение с камеры для последующего преобразования к эквидистантной панораме в виде двух проекций изображения, в которых прямые линии заменяются на дугообразные кривые, а сами проекции представляют собой круги (рис. 1, 2) [1].

В настоящее время существует реализация алгоритма преобразования изображения к сферической панораме в библиотеке компьютерного зрения OpenCV для языка C++ и в ее аналоге BoofCV для языка Java. Обе библиотеки имеют открытый исходный код. Для реализации преобразования с использованием этих библиотек необходимо совершать калибровку камеры с использованием соответствующих методов используемой библиотеки.

При калибровке камеры вычисляются параметры линз для последующего использования в алгоритме преобразования [2, 3].

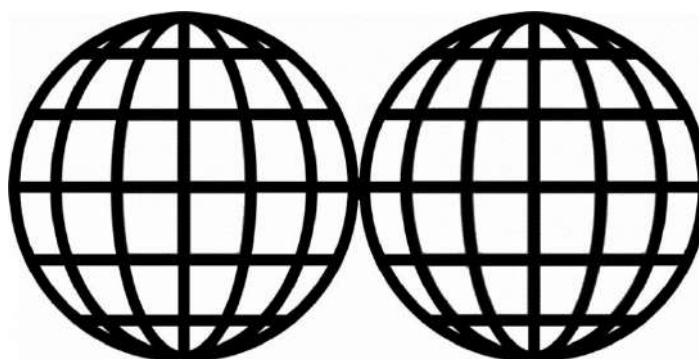


Рисунок 1. Координатная сетка исходного изображения

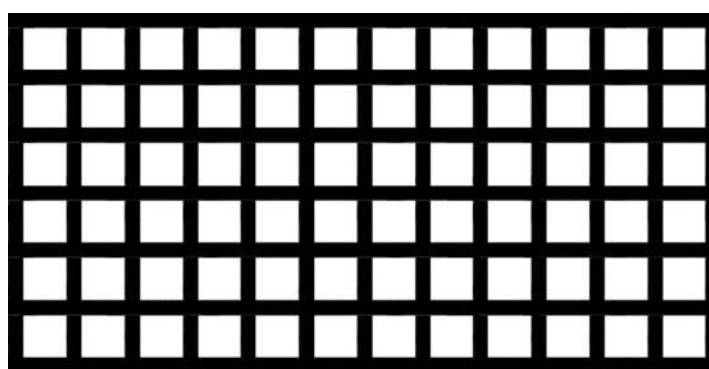


Рисунок 2. Координатная сетка эквидистантной панорамы

Главным недостатком библиотек являются неточности калибровки, что непосредственно влияет на качество получаемых эквидистантных панорам при работе алгоритма преобразования [4].

Наиболее популярными программными средствами для обработки и преобразования изображений к формату сферических панорам являются Kolor Autopano Giga и Gear 360 ActionDirector. Из недостатков этих программ следует выделить:

- большое количество ошибок преобразования в области наложения изображений для программы Gear 360 ActionDirector;
- сложный пользовательский интерфейс и, как следствие, высокое время, затрачиваемое пользователем на ручное создание модели преобразования изображения для программы Kolor Autopano Giga.

Содержание исследования

На основании особенностей исследованных программных решений и в соответствии с потребностями пользователей были составлены функциональные и нефункциональные требования к программному решению: удобный и ненагруженный пользовательский интерфейс, возможность автономной обработки изображения без излишних настроек программного средства.

Для реализации программного решения был выбран объектно-ориентированный кроссплатформенный язык Java 8, который имеет большие возможности для решения различных задач собственными средствами и обеспечивает необходимую поддержку сторонних библиотек [5]. Для реализации методов преобразования изображения была выбрана библиотека VoofCV.

Разработанная программа содержит удобный пользовательский интерфейс и возможность обработки большого количества фотографий без дополнительных настроек.

Для каждого исследованного коммерческого решения и собственного реализованного решения был проведен ряд экспериментов по обработке изображения, в ходе которых измерялось время, затрачиваемое пользователем на настройку программы, и время работы алгоритма – время рендера. Средние значения измеренных величин представлены в таблице 1.

Таблица 1. Сравнение временных характеристик работы программ

	Gear 360 ActionDirector		Kolor Autopano Giga		Разработанное приложение	
Количество изображений, файлов	1	20	1	20	1	20
Среднее время, затраченное пользователем, с	25,3 ± 0,5	25,3 ± 0,6	127,25 ± 30	2500,0 ± 60	16,05 ± 2,8	16,04 ± 2,8
Среднее время рендера, с	3,73 ± 0,12	74,3 ± 0,15	3,37 ± 0,06	70,3 ± 0,15	5,7 ± 0,04	78,0 ± 0,6

Заключение

Рассмотрены существующие подходы к автоматизации обработки изображения. Разработана программа, автоматизирующая многие действия пользователя и сокращающая время, затрачиваемое пользователем на обработку изображения. Преобразование изображения осуществляется с приемлемым для пользователя качеством.

Разработанная программа имеет недостатки в работе алгоритма, которые планируется устранить в ходе дальнейших исследований следующими способами:

- использование других реализаций алгоритма калибровки или доработка существующей;
- введение многопоточности в работу алгоритма преобразования изображения.

ЛИТЕРАТУРА

1. *Л. Шапиро, Дж. Стокман. Компьютерное зрение.* / Л. Шапиро, Дж. Стокман – М.: Бином. Лаборатория знаний, 2006. – 725 с.
2. *Bradski, G. Learning OpenCV – Computer Vision with the OpenCV Library* / G. Bradski, A. Kaehler. – O'Reilly Media, 2008. – 580 с.
3. *Эккель, Б. Философия Java* / Б. Эккель. – М.: Питер, 2016. – 809 с.
4. *Глаголев В.М. Описание и программное устранение дисторсии объективов* // Известия тувского государственного университета. Технические Науки, 2017.
5. *Эккель, Б. Философия Java* / Б. Эккель. – М.: Питер, 2016. – 809 с.

Сведения об авторе

Ирина Андреевна Стоколяс — студент, Университет ИТМО,
Факультет программной инженерии и компьютерной техники.
Email: ira.stokolias@gmail.com

Ссылка для цитирования: *Стоколяс И.А.* Подходы к автоматизации обработки изображения формата 360 градусов // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 57—60.

**APPROACHES TO AUTOMATION OF IMAGE PROCESSING
OF THE FORMAT OF 360 DEGREES****I.A. Stokolias***ITMO University, 197101, St. Petersburg, Russia*

The progress of photographic and video equipment production technologies for shooting widescreen images revealed the necessity for processing spherical panoramas automation. The main software tools for processing spherical panoramas were studied and their advantages and disadvantages were determined. Based on those results a software tool was developed that considers user's requirements and needs when working with the program. Experimental results indicate that the developed program has successfully automated user's work.

Keywords: BoofCV, OpenCV spherical panorama, equidistant panorama, image processing, camera calibration, graphical user interface, automation

Irina A. Stokolias	—	Data on author student; ITMO University, Faculty of Software Engineering and Computer Systems; E-mail: ira.stokolias@gmail.com
---------------------------	---	--

For citation: *Stokolias I.A.* Approaches to automation of image processing of the format of 360 degrees // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 57—60 (in Russian).

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ ПРОГРАММНАЯ ИНЖЕНЕРИЯ

УДК 004.056.2

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ ТЕХНОЛОГИИ БЛОКЧЕЙН В ОБЛАСТИ ЗДРАВООХРАНЕНИЯ

Г.Б. Аукен, А.В. Белозубов

Университет ИТМО, 197101, Санкт-Петербург, Россия

В этой статье ставится вопрос о том, должны ли информационные системы быть централизованными или децентрализованными для обеспечения большей поддержки различных отраслей. Целью работы является исследование принципов работы технологии Блокчейн и многомерных баз данных. Внедрение новой системы поможет исключить многократное прохождение одного и того же анализа. Кроме того, медицинские организации смогут воспользоваться этой информацией для того, чтобы поставить правильный диагноз или назначить процедуры. Помимо этого, каждый человек получит личный кабинет и доступ к интересующей его информации. Использование технологии Блокчейн может увеличить эффективность работы медицинских учреждений. Блокчейн также похож на базу данных, в которой хранится информация, однако основное различие заключается в том, что данные расположены в сети персональных компьютеров, называемых узлами, где нет центральной организации, такой как правительство или банк, контролирующей данные.

Ключевые слова: информация, технология, базы данных, программирования, хранилища данных

Введение

Эта статья касается развития организационной инфраструктуры информационных технологий (ИТ) в учреждениях с высокой степенью независимости на уровне медицинских учреждений в разработке ресурсов информационных технологий.

Многомерные базы данных отличаются от реляционных прежде всего трехмерностью — поддержкой неограниченного числа значений в поле, и находят свое применение там, где необходима эффективная и простая работа с большими массивами символьной информации. В многомерных СУБД данные организованы в виде упорядоченных многомерных массивов, удовлетворяющих требованиям защиты от несанкционированного доступа в организации. Они обеспечивают более быструю реакцию на запросы данных за счет того, что обращения поступают к относительно небольшим блокам данных, необходимых для конкретной группы пользователей. Для достижения сравнимой производительности реляционные системы требуют тщательной проработки схемы базы данных, определения способов индексации и специальной настройки [1].

Необходима всеобъемлющая структура хранилища данных, которая включает в себя визуализацию и информацию, не связанную с изображениями, в поддержке управления болезнями и исследований. Авторы предлагают такую структуру, описывают общие принципы проектирования и архитектуру системы и иллюстрируют мультимодальную

систему хранения данных для нейровизуализации, внедренную для медицинских исследований. Часто информация разбросана по нескольким объектам, и иногда она недоступна, когда она необходима больше всего – это та ситуация, которая разыгрывается каждый день вокруг нас [2]. И эту проблему медицины может решить технология Блокчейн.

Обзор предметной области

В качестве примера, можно привести ситуацию, когда врач видит пациента или записывает новый рецепт, пациент соглашается иметь ссылку или «указатель», добавленный в блок-цепь, - децентрализованный цифровой регистр, такой как базовый регистр. Вместо того, чтобы осуществлять платежи, этот блок-код будет записывать критическую медицинскую информацию в практически неподкупную криптографическую базу данных, поддерживаемую сетью компьютеров, которая доступна для всех, кто работает с программным обеспечением. Каждый указатель, который вводит врач, станет частью записи пациента, независимо от того, какую электронную систему использовал врач, поэтому любой врач может использовать его, не беспокоясь о проблемах несовместимости [3].

Технологи и специалисты в области здравоохранения по всему миру видят технологию блокчейн, как способ упорядочить совместное использование медицинских записей безопасным способом, защитить конфиденциальные данные от хакеров и дать пациентам больше контроля над своей информацией [4, 5].

Технология блокчейн предлагает нам:

1. Отсутствие центрального сервера;
2. Быстрые и точные транзакции;
3. Наличие полной копии базы;
4. Шифрование информации.

Содержание исследования

Программирование децентрализованных информационных систем дает нам возможность увеличить скорость доступа к данным. Технология блокчейн позволяет избежать многих сложностей, возникающих при работе с базами данных. Многомерные базы данных — технология, которая длительное время воспринималась как новинка, — сегодня является решением, которое предлагает не только высокую производительность и простоту использования, но и обеспечивает возможности, необходимые для разработки, расширения и быстрого развертывания бизнес-приложений при сокращении ИТ-затрат. Системы на основе многомерных баз данных идеально подходят для потребностей как для рынков среднего и малого бизнеса (SMB), так и крупных предприятий.

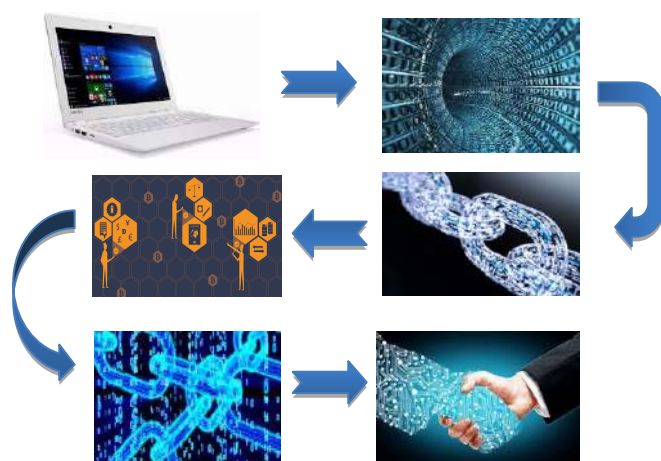


Рисунок. Принцип работы технологии блокчейн

Как мы видим на рисунке, принцип работы технологии блокчейн заключается в следующем:

1. Пациент пересылает свои данные участникам
2. Транзакция передается в сеть
3. Блоки рассылаются для проверки всем участникам системы
4. Каждый участник записывает блок в свой экземпляр БД
5. Блок попадает в «цепочку блоков» которая содержит информацию обо всех транзакциях
6. Транзакция завершается

Заключение

Промежуточные результаты исследования показывают, что использование технологии блокчейн может увеличить эффективность работы медицинских учреждений. Блокчейн также похож на базу данных, в которой хранится информация, однако основное различие заключается в том, что данные расположены в сети персональных компьютеров, называемых узлами, где нет центральной организации, такой как правительство или банк, контролирующей данные.

ЛИТЕРАТУРА

1. *Abellán P., Tos T D., GR AU S., Puig A.* Иллюстративная визуализация мультимодальных наборов данных на основе регионов // *Computerized Medical Imaging and Graphics* 37, 4 (2013), 263–271. Elsevier, Amsterdam.
2. A Survey on Multimodal Medical Data Visualization (PDF Download Available) [Электронный ресурс] <https://www.researchgate.net/publication/320600348> A_Survey_on_Multimodal_Medical_Data_Visualization [accessed Mar 15 2018].
3. *Bailey D.L., Wil Lowson K.P.* Основанный на фактических данных обзор количественной визуализации и потенциальных клинических применений // *Journal of Nuclear Medicine* 54, 1 (2013), 83–89. SNMMI. A Survey on Multimodal Medical Data Visualization (PDF Download Available) [Электронный ресурс]: <https://www.researchgate.net/publication/320600348> A_Survey_on_Multimodal_Medical_Data_Visualization [accessed May 15 2018].
4. *Аукен Г.Б.* Современные информационные технологии в медицине // 19 Международные научные чтения – 2017. – С20-24
5. *Белозубов А.В.* Место многомерных и мультимодальных данных в информационной структуре медицины // Материалы международной научно-практической конференции «Современные тенденции в образовании и науке: состояние и перспективы» - 16 марта 2018 – КЭУК. - с.191-193

Сведения об авторе

- Гулим Аукен** — аспирант; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: gulim.auken@inbox.ru
- Александр Владимирович Белозубов** — доцент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: belozubov@corp.ifmo.ru

Ссылка для цитирования: Аукен Г., Булозубов А.В. Программируемые децентрализованные информационные системы // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 61—64.

PROGRAMMABLE DECENTRALIZED INFORMATION SYSTEMS

G. B. Auken, A.V. Belozubov

ITMO University, 197101, St. Petersburg, Russia

This article asks whether information systems should be centralized or decentralized to provide greater support for various industries. The aim of the work is to study the principles of operation of the Blockchain technology and multidimensional databases. The introduction of a new system will help eliminate the repeated passage of the same analysis. In addition, medical organizations will be able to use this information in order to make a correct diagnosis or prescribe procedures. In addition, each person will receive a personal account and access to information of interest to him. The use of Blockchain technology can increase the efficiency of medical institutions. The blockchain is also similar to the database in which information is stored, but the main difference is that the data is located on a network of personal computers, called nodes, where there is no central organization, such as a government or a bank controlling data

Keywords: information, technology, databases, programming, data warehouses

Data on author

- Gulim Auken** — graduate student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: gulim.auken@inbox.ru
- Alexandr V. Belozubov** — docent; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: belozubov@corp.ifmo.ru

For citation: Auken G., Belozubov A.V. Programmable decentralized information systems // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 61—64 (in Russian).

УДК 004.052.3

МОДЕЛЬ НАДЕЖНОСТИ КЛАСТЕРА С ПОДДЕРЖКОЙ НЕПРЕРЫВНОСТИ ВЫЧИСЛИТЕЛЬНОГО ПРОЦЕССА ПРИ МИГРАЦИИ ВИРТУАЛЬНЫХ МАШИН

А.Н. Деркач¹, С.М. Алексанков²

1 - Университет ИТМО, 197101, Санкт-Петербург, Россия

2 - НИИ Масштаб, 194100, Санкт-Петербург, Россия

Рассмотрена марковская модель надежности отказоустойчивого кластера с использованием технологий виртуализации, обеспечивающих непрерывность вычислительного процесса в случае сбоя физических ресурсов серверов и невозможности восстановления после прерывания вычислительного процесса. В модели учтено влияние на надежность условий восстановления, реализованных или не реализованных на определенных этапах деградации системы.

Ключевые слова: виртуализация, надежность, отказоустойчивость, кластер, миграция виртуальных машин

Введение

Наиболее важными в серверных вычислительных системах, в особенности критических системах реального времени, являются параметры надежности и отказоустойчивости, которые имеют значение для обеспечения непрерывного доступа пользователя к необходимым операциям. Достижению высоких и стабильных показателей производительности, надежности, отказоустойчивости [1-4] и безопасности вычислительных компьютерных систем способствует применение актуальных технологий кластеризации и виртуализации [5-6]. Поддержание данных параметров на высоком уровне и создание высокой отказоустойчивости вычислительных процессов [7, 8] при виртуализации обеспечивается миграцией виртуальных ресурсов между физическими серверами, являющимися узлами кластера. При выходе из строя физических ресурсов миграция виртуальных машин обеспечит ускорение процесса реконфигурации.

Обзор предметной области

Одним из эффективных способов достижения требуемой степени отказоустойчивости вычислительных процессов во время виртуализации является миграция виртуальных ресурсов между физическими узлами (серверами) кластера [9]. Рассмотрена проблема повышения надежности на основе виртуализации с применением технологии «Отказоустойчивость». Технология «Отказоустойчивость», обеспечивает непрерывность вычислительного процесса (сервиса) в кластере после отказа одного физического сервера при поддержке двух копии ВМ в оперативной памяти, расположенные на разных физических серверах, так чтобы в случае отказа одного из них продолжить работу на втором. Для рассматриваемой организации вычислительного процесса во время функционирования ВМ на одном из серверов, на втором должна поддерживаться актуальная копия оперативной памяти [10-11] активной ВМ. При этом образы виртуальных дисков ВМ должны храниться на выделенном или распределенном хранилище данных с синхронной репликацией данных.

Цель работы – исследование возможности повышения функциональной надежности кластерных систем с миграцией виртуальных машин при недопущении нарушения непрерывности вычислительного процесса.

Под функциональной надежностью будем понимать способность систем выполнять функции с учетом не только работоспособности требуемых для их выполнения ресурсов, но и обеспечения условий реализации выполняемых процессов хранения, передачи и обработки данных. В качестве условий требуемого функционирования может выдвигаться условие обеспечения непрерывности вычислительного процесса при недопустимости перерывов работы резервированной системы на время восстановления. Таким образом в рассматриваемых системах восстановление возможно только при его совмещении с реализацией требуемых функций не отказавшими узлами. В случае невозможности реконфигурации с активацией необходимого числа работоспособных ресурсов система переходит в состояние невозстанавливаемого отказа.

Содержание исследования

Объект исследования – отказоустойчивый двухмашинный кластер, объединяющий физические узлы при помощи коммутатора. Кластер содержит два сервера с локальным подключенными устройства хранения, Связи между серверами осуществляется через коммутатор. Система хранения данных представлена в виде локального хранилища для каждого физического узла в виде жесткого диска. Технология отказоустойчивости поддерживает синхронную репликацию данных между локальными хранилищами.

Для рассматриваемого двухмашинного кластера в работе [10-11] построена марковская модель на основе которой определены стационарный и нестационарный коэффициенты готовности.

Следует заметить, что использование коэффициента готовности позволяет оценить только готовность системы к выполнению требуемых функций, но не характеризует кластер как сложную систему, способную находиться в множестве работоспособных состояний с разной эффективностью функционирования. В связи с этим для рассматриваемой кластерной системы в предлагаемой статье предлагается использовать коэффициент сохранения эффективности.

$$K = P_0 + \sum_{i \in S} \frac{E_i}{E_0} P_i,$$

Где P_i вероятность i -го работоспособного состояния, P_0 – вероятность исходного состояния с исправностью всех узлов кластера, S – множество работоспособных состояний кластера с отказами отдельных узлов. E_0 и E_i – показатели эффективности исходного и i -го состояний. Рассматривая в качестве показателей эффективности среднее время пребывания запросов коэффициент сохранения эффективности определим как

$$T_1 = P_0 + \sum_{i \in S} \frac{T_0}{T_i} P_i = P_0 + \frac{T_0}{T_1} \sum_{i \in S} P_i,$$

В исходном состоянии систему представим в виде двух одноканальных систем массового обслуживания типа М/М/1. А для работоспособных состояний с отказами отдельных узлов одной одноканальных систем массового обслуживания типа М/М/1. Время пребывания для исходного исправного состояния и работоспособные состояния с отказами узлов найдем соответственно как

$$T_0 = \frac{v}{1 - 0,5\Lambda v}, \quad T_1 = \frac{v}{1 - \Lambda v}$$

Для оценки вероятностей работоспособных состояний кластера можно воспользоваться системой уравнений из [11]

$$\begin{cases}
0 = -(2\lambda_0 + \lambda_2 + 2\lambda_1)P_0(t) + \mu_3 P_4(t), \\
0 = -(\lambda_1 + \lambda_0 + \mu_0)P_1(t) + \lambda_0 P_4(t) + 2\lambda_0 P_0(t) + \mu_4 P_{11}(t) + \mu_0 P_6(t), \\
0 = -(\lambda_1 + \lambda_0 + \mu_2)P_2(t) + \mu_1 P_7(t) + \mu_0 P_8(t) + \lambda_2 P_4(t) + \lambda_2 P_0(t), \\
0 = -(\lambda_1 + \lambda_0 + \mu_1)P_3(t) + \mu_1 P_9(t) + \mu_0 P_{10}(t) + \mu_4 P_{12}(t) + \mu_0 P_5(t) + \lambda_1 P_4(t) + 2\lambda_1 P_0(t), \\
0 = -(\lambda_1 + \lambda_0 + \lambda_2 + \mu_3 + \lambda_1 + \lambda_0)P_4(t) + \mu_0 P_1(t) + \mu_2 P_2(t) + \mu_1 P_3(t), \\
0 = -\mu_0 P_5(t) + \lambda_1 P_1(t) + \lambda_1 P_{11}(t), \\
0 = -\mu_0 P_6(t) + \lambda_0 P_1(t) + \lambda_1 P_0(t), \\
0 = -\mu_1 P_7(t) + \lambda_1 P_2(t), \\
0 = -\mu_0 P_8(t) + \lambda_0 P_2(t), \\
0 = -\mu_1 P_9(t) + \lambda_1 P_3(t) + \lambda_1 P_{12}(t), \\
0 = -\mu_0 P_{10}(t) + \lambda_0 P_3(t) + \lambda_0 P_{12}(t), \\
0 = -\mu_4 P_{11}(t) + \lambda_0 P_4(t), \\
0 = -\mu_4 P_{12}(t) + \lambda_1 P_4(t),
\end{cases}$$

где обозначены интенсивности отказов λ_0 и восстановлений μ_0 сервера; диска λ_1 , μ_1 ; коммутатора λ_2 , μ_2 . Интенсивность восстановления (синхронизации системы распределенного хранилища) – μ_3 , включающего занесении актуальной реплики данных на восстановленный диск. Интенсивность восстановления виртуальной машины после автоматического перезапуска μ_4 , включающего запуск ВМ на резервном сервере и загрузку на нем приложения пользователя.

Заключение

Таким образом, для кластерных систем с миграцией виртуальных машин при недопущении нарушения непрерывности вычислительного процесса рассмотрена оценка коэффициента сохранения эффективности с учетом возможности нахождения системы в работоспособных состояниях с отказами отдельных узлов кластера.

ЛИТЕРАТУРА

1. Kopetz H. *Real-Time Systems: Design Principles for Distributed Embedded Applications*. Springer, pp. 396, 2011.
2. Sorin D. *Fault Tolerant Computer Architecture*. Morgan & Claypool 2009. 103 p.
3. Dudin, A. N., Sun, B. A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers // *Automatic Control and Computer Sciences*, 2009, No. 5, pp. 32–44.
4. Aliev T.I., Rebezova M.I., Russ A.A. Statistical Methods for Monitoring Travel Agencies // *Automatic Control and Computer Sciences* - 2015, Vol. 49, No. 6, pp. 321–327
5. Богатырев В.А. Оптимальное резервирование системы разнородных серверов // *Приборы и системы. Управление, контроль, диагностика* -2007. - № 12. - С. 30-36
6. Богатырев В.А., Богатырев А.В., Голубев И.Ю., Богатырев С.В. Оптимизация распределения запросов между кластерами отказоустойчивой вычислительной системы // *Научно-технический вестник информационных технологий, механики и оптики* -2013. - № 3(85). - С. 77-82
7. Богатырев В.А., Богатырев С.В., Богатырев А.В. Надежность кластерных вычислительных систем с дублированными связями серверов и устройств хранения // *Информационные технологии* -2013. - № 2. - С. 27-32

8. Богатырев В.А., Богатырев С.В. Надежность мультикластерных систем с перераспределением потоков запросов // Известия высших учебных заведений. Приборостроение -2017. - Т. 60. - № 2. - С. 171-177
9. Алексанков С.М. Модель процесса динамической миграции с копированием данных после остановки виртуальных машин // Изв. вузов. Приборостроение. 2016. Т. 59. № 5. С. 173–178
10. Алексанков С.М., Богатырев В.А., Деркач А.Н. Надежность отказоустойчивого кластера с учетом миграции виртуальных машин и оперативности дисциплин восстановления // Вестник компьютерных и информационных технологий. 2018. №9. С. 49-56.
11. Алексанков С.М., Богатырев В.А., Деркач А.Н. Модель надежности отказоустойчивого кластера с миграцией виртуальных машин // Программные продукты и системы -2019. Т32- № 1. С 561-566.

Сведения об авторе

- | | | |
|-------------------------------------|---|--|
| Алексей Николаевич Деркач | — | аспирант; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: alexitmo1@gmail.com |
| Сергей Михайлович Алексанков | — | инженер; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: Aleksankov.sergey@gmail.com |

Ссылка для цитирования: Деркач А.Н., Алексанков С.М. Модель надежности кластера с поддержкой непрерывности вычислительного процесса при миграции виртуальных машин // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 65—68.

MODEL OF CLUSTER RELIABILITY WITH SUPPORT OF CONTINUITY OF COMPUTATIONAL PROCESS IN THE MIGRATION OF VIRTUAL MACHINES

A.N. Derkach¹, S.M. Aleksankov²,

1 - ITMO University, 197101, St. Petersburg, Russia

2 - Research Institute Mashtab, 194100, St. Petersburg, Russia

A Markov model of reliability of a fault-tolerant cluster has been considered, using virtualization technologies that ensure the continuity of the computational process in the event of a failure of the servers' physical resources and the impossibility of recovering from the interruption of the computational process. The model takes into account the impact on the reliability of restoration conditions, whether implemented or not implemented at certain stages of system degradation. The criterion for assessing the reliability of the study is the assessment of the efficiency's conservation.

Keywords: virtualization, reliability, fault tolerance, cluster migration of virtual machines

Data on author

- | | | |
|----------------------------|---|--|
| Alexey N. Derkach | — | graduate student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: alexitmo1@gmail.com |
| Sergey M. Alexankov | — | engineer; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: Aleksankov.sergey@gmail.com |

For citation: Derkach A.N., Alexankov S.M. Model of cluster reliability with support of continuity of computational process in the migration of virtual machines // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 65—68 (in Russian).

УДК 004.4

ПОДХОД К ПРОГНОЗИРОВАНИЮ ПОВЕДЕНИЯ СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Л.С. Лисицына, С.А. Орешин

Университет ИТМО, 197101, Санкт-Петербург, Россия

В настоящее время сетевые информационные системы накапливают большой объем статистических данных, что открывает большие возможности для моделирования и прогнозирования поведения сложных систем и разработки новых рекомендательных сервисов в них. Для решения таких задач успешно используются разнообразные методы машинного обучения, которые постоянно совершенствуются и развиваются. Данная статья посвящена разработке и исследованию нового подхода прогнозирования поведения сложных систем на основе применения различных алгоритмов градиентного бустинга. Особенностью данного подхода является использование 6-ти различных моделей для построения общей регрессионной модели, предсказывающей спрос на такси в заданном районе на 6 часов, начиная с текущего момента времени. В статье приведен анализ прогнозирования временных рядов, который подтверждает эффективность предложенного подхода в сравнении с классическими методами.

Ключевые слова: машинное обучение, авторегрессия, анализ данных, прогнозирование

Введение

В настоящее время растет спрос на прогнозирование поведения сложных систем из-за большого количества данных, которые позволяют накапливать и обрабатывать современное оборудование. Сложная система – это система, состоящая из множества различных объектов и событий, итоговое взаимодействие которых невозможно описать алгоритмически.

Для формализации задачи будем фиксировать интересующий нас параметр и рассматривать его значения через равные промежутки времени. Полученный график принято называть временным рядом. В статье рассматриваются классические подходы к прогнозированию временных рядов, а также приводится задача, где они не являются неэффективными. Целью исследования является разработка нового подхода к прогнозированию временных рядов на основе алгоритмов градиентного бустинга и доказательства его эффективности.

Обзор предметной области

Основная задача, рассматриваемая в статье – это предсказание временных рядов. Общепринятое решение для задач предсказания временных рядов – это использование модели авторегрессии скользящего среднего (класс моделей ARMA) [1]. Однако, данный подход имеет ряд недостатков. Для применения моделей ARMA для предсказания временных рядов с сезонностью необходимо добавить сезонную компоненту и приводить ряд к стационарному перед применением модели [2]. В данной статье рассматривается задача, в которой использование классического подхода оказывается неэффективным из-за большого количества временных рядов и невозможности использовать внешних признаков.

Содержание исследования

Для предсказания спроса в различных районах Нью-Йорка предложено разбить карту города на 2500 равных квадратов, и агрегировать вызовы такси по этим областям. Будем рассматривать области со средним спросом более 5 вызовов в час, таким образом выделим 102 клетки, для которых будем строить прогноз. Полученный результат приведен на рисунке 1.

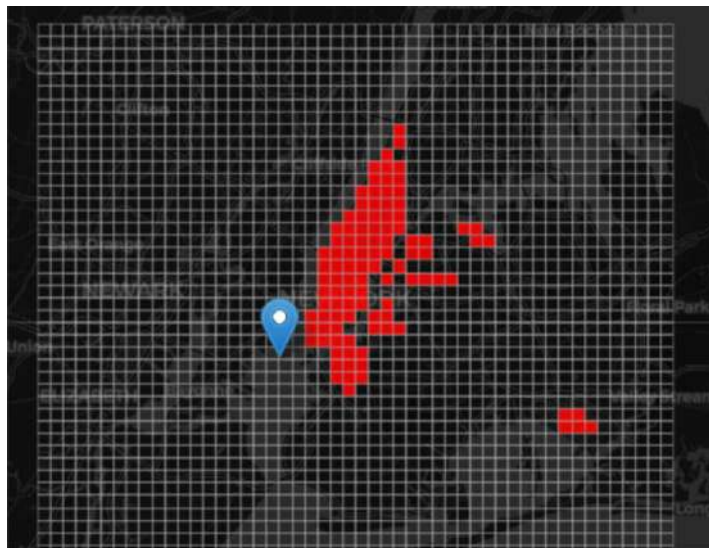


Рисунок 1. Полученное пространство поиска

Для построение первой модели возьмем ряд, в котором расположен небоскреб Empire State Building. Для применения модели SARIMA приведем ряд с стационарному с помощью сезонного и почасового дифференцирования и логарифмирования [2]. Полученный результат предсказания приведен на рисунке 2. По полученному результату видно, что модель учитывает сезонность и цикличность при построении прогноза. Однако, для прогнозирования всех полученных регионов, необходимо вручную обрабатывать 102 ряда, а также перебирать большое количество моделей, чтобы выбрать оптимальную по информационному критерию Акаике [3]. По примерным расчетам, полный прогноз с применением данного подхода займет около 900 часов, поэтому подход с использованием обычных моделей SARIMA не считается оптимальным.

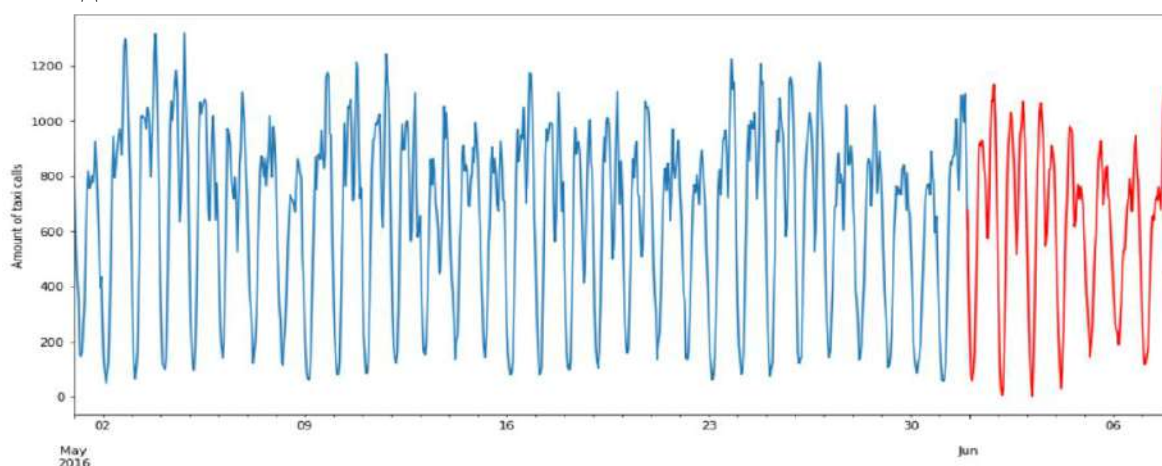


Рисунок 2. Прогноз временного ряда моделью SARIMA на неделю

Для применения моделей класса SARIMA кластеризируем ряды, и будем вручную обрабатывать ряд, полученный усреднением всех рядов в каждом кластере, а затем применим обученную на этом ряде модель, на все остальные. Результат кластеризации алгоритмом k-means [4] с количеством кластеров 5.

Приведенный выше подход имеет ряд недостатков для решения данной задачи. В статье предлагается использовать подход, с применением 6 различных моделей градиентного бустинга [5] для предсказания на различное количество часов в будущем, обученных на различных значениях скользящего среднего, лагах, полученных применением аппроксимации с помощью рядов Фурье, а также различными географическими и временными признаками – количество вызовов из соседних клеток, государственные праздники и др. Обучение моделей и построение прогноза данным подходом заняло около 12 часов.

В таблице 1 представлен сравнительный анализ моделей ARIMA, обученных после кластеризации и 6 итоговых моделей. В качестве метрики качества использовалась нормированная абсолютная ошибка по всем рядам по предсказанию на 6 часов в будущее:

$$Q = \frac{1}{R \times 739 \times 6} \sum_{r=1}^R \sum_{T=2016.05.31\ 23:00}^{2016.06.30\ 17:00} \sum_{i=1}^6 |\hat{y}_{T|T+i}^r - y_{T+i}^r|.$$

Таблица 1. Сравнительный анализ различных подходов

	Предложенный подход	Подход с кластеризацией и SARIMA
Q	17.7412	31.1128
Общее время вычислений (в часах)	12	30
Возможность учитывать внешние признаки	Да	Нет
Необходимость предобработки рядов	Нет	Да

Заключение

Представленный в данной статье подход имеет ряд преимуществ по сравнению с классическим подходом прогнозирования временных рядов с использованием моделей SARIMA. Основные преимущества подхода – это сокращение времени вычислений и возможность применения внешних признаков. Данный подход применим к предсказанию поведения различных сложных систем

ЛИТЕРАТУРА

1. Mills, Terence C. Time Series Techniques for Economists // Cambridge University Press. ISBN 0-521-34339-9. 1990.
2. Hamilton, James. Time Series Analysis, Princeton University Press, ISBN 0-691-04289-6. 1994.
3. Akaike, H. Prediction and entropy // A Celebration of Statistics. Springer. 1985. P. 1–24.
4. Bailey, Ken. "Numerical Taxonomy and Cluster Analysis" // Typologies and Taxonomies. ISBN 9780803952591. 1994. P. 34.
5. Hastie, T.; Tibshirani, R.; Friedman, J. H. 10. Boosting and Additive Trees // The Elements of Statistical Learning (2nd ed.). New York. Springer. ISBN 0-387-84857-6. 2009. P. 337–384.

Сведения об авторе

- Святослав Анатольевич Орешин** — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: aqice26@gmail.com
- Любовь Сергеевна Лисицына** — профессор; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: lisizina@corp.ifmo.ru

Ссылка для цитирования: Орешин С.А., Лисицына Л.С. Подход к прогнозированию поведения сложных систем на основе методов машинного обучения // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 70—73.

APPROACH TO FORECASTING THE BEHAVIOR OF COMPLEX SYSTEMS BASED ON MACHINE LEARNING METHODS**L. S. Lisitsyna, S. A. Oreshin***ITMO University, 197101, St. Petersburg, Russia*

Currently, network information systems accumulate a large amount of statistical data, which opens up great opportunities for modeling and forecasting the behavior of complex systems and the development of new Advisory services in them. To solve such problems, a variety of machine learning methods are successfully used, which are constantly being improved and developed. This article is devoted to the development and research of a new approach to predicting the behavior of complex systems based on the use of various gradient boosting algorithms. The peculiarity of this approach is the use of 6 different models to build a General regression model that predicts the demand for a taxi in a given area for 6 hours, starting from the current time. The article presents an analysis of time series forecasting, which confirms the effectiveness of the proposed approach in comparison with classical methods.

Keywords: Machine learning, autoregression, data analysis, forecasting

Data on author

- Svjatoslav A. Oreshin** — student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: aqice26@gmail.com
- Lubov S. Lisizina** — professor; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: lisizina@corp.ifmo.ru

For citation: Oreshin S.A., Lisizina L.S. Approach to forecasting the behavior of complex systems based on machine learning methods // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 69—72 (in Russian).

УДК 004.37

ИССЛЕДОВАНИЕ ЗАВИСИМОСТИ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ ОТ ВРЕМЕНИ ИЗУЧЕНИЯ ТЕОРЕТИЧЕСКОГО МАТЕРИАЛА СТУДЕНТАМИ НА ОТКРЫТОМ НАБОРЕ ДАННЫХ

Н.Д. Матросова, Д.Г. Штенников

Университет ИТМО, 197101, Санкт-Петербург, Россия

Была проверена гипотеза о зависимости итоговой оценки онлайн-курса от временных характеристик изучения лекционного материала. Исследования были проведены на открытом наборе данных, предоставленных репозиторию машинного обучения UCI.

Ключевые слова: языки описания аппаратуры, высокоуровневый синтез, ПЛИС, Haskell, поиск строк

Введение

Данное исследование является предварительным этапом для построения рекомендательного алгоритма, а именно исследованием возможности использования открытых наборов данных для поставленных задач. Как известно, пользователи по-разному взаимодействуют с информационными системами (ИС). От последовательности и характера взаимодействий может зависеть результат пользователя в данной системе. Такой «след активности» пользователя может содержать различную информацию, в зависимости от настроек ИС.

Для исследования рассматривались информационные образовательные системы, так как интерес к использованию онлайн-курсов подводит исследователей к вопросам о влиянии цифровых моделей преподавания на студентов и возможности адаптации учебных дистанционных курсов под индивидуальности потребности и особенности студентов. Надо также учитывать, что распространенность сети Интернет вызвала рост интереса пользователей к самообразованию, чаще всего выраженной в просмотре видео-контента, обращениям к материалам цифровых библиотек и посещением онлайн-курсов [1].

Основной составляющей онлайн-курсов являются лекционные материалы и тестирования, поэтому авторы предположили наличие зависимости продолжительности изучения лекционного материала и итоговой оценки за тестирование.

Обзор предметной области

В 2017 году сервис CourseBug провел исследование, посвященное анализу и сравнению пяти платформ, реализующих MOOC: «Coursera», «EdX», «Открытое образование», «Универсариум» и Лекториум». Основной целью исследования было выявление количества людей, готовых обучаться или уже обучающихся онлайн. По официальным данным суммарное количество посещений со всех платформ в период с сентября 2014 года по август 2017 составило 60,7 миллионов пользователей [2,3]. Данное исследование демонстрирует интерес пользователей к использованию онлайн-курсов. Один из анализов активностей обучающихся (для изучения проблем и вопросов адаптации учебных материалов) был представлен в 2016 года на блоге платформы «Udemy» в виде инфографики, посвященной образовательным трендам года [4]. Анализ содержал информацию о возрастных категориях пользователей, самых популярных авторов и курсов, распределение количества человек по

городам. Отличительной особенностью анализа были сведения о днях неделях, в которых чаще всего и реже всего обучаются пользователи (среда и воскресенье, соответственно).

Работая с курсом, пользователь оставляет за собой достаточно широкий «след активности». Например, в LMS Moodle это демонстрируется с помощью логов событий, в которые собирается информация о времени обращения к контенту, последовательность работы с элементами и т.п., а при необходимых настройках может отображаться и более расширенная статистика [5].

Содержание исследования

Для исследования в качестве исходных данных был использован набор данных Educational Process Mining (EPM): A Learning Analytics Data Set Data Set, доступный в репозитории машинного обучения UCI [6,7]. Этот набор содержит данные обучения группы студентов. Набор содержит информацию с шести сессий по 115 пользователей с учетом 13 характеристик. В конце всех шести сессий студентам было предложено сдать экзамен, на который имелось две попытки. Общее количество логов за шесть сессий - 230318. Характеристика «время простоя» в представленных данных у некоторых студентов обладала отрицательной величиной, про которую авторы-представители набора данных не дали каких-либо пояснений в описании. Поэтому данную характеристику пришлось исключить из исследования.

На первом этапе авторы отобрали данные, содержащиеся в столбцах «ID студента», «Активный элемент», «Время начала» и «Время окончания» для выявления информационных элементов. В работе были использованы следующие дополнительные характеристики - «Среднее время», «Суммарное время» и «Медианное время» работы студента с информационным элементом, вычисленным самостоятельно авторами исследования. Все три характеристики можно отнести к интервальной шкале. Для определения корреляционных зависимостей был использован коэффициент корреляции Пирсона:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

По результатам проведенных расчетов коэффициент корреляции Пирсона для пары характеристик «Медианное время» и «Итоговая отметка» равен 0.17; для пары «Суммарное время» и «Итоговая отметка» коэффициент равен 0.22; для последней пары «Среднее время» и «Итоговая отметка» коэффициент равен 0.18 – во всех трех случаях это означает очень слабую или слабую корреляцию. Это подтверждается графиками рассеивания, представленных на рисунке 1.

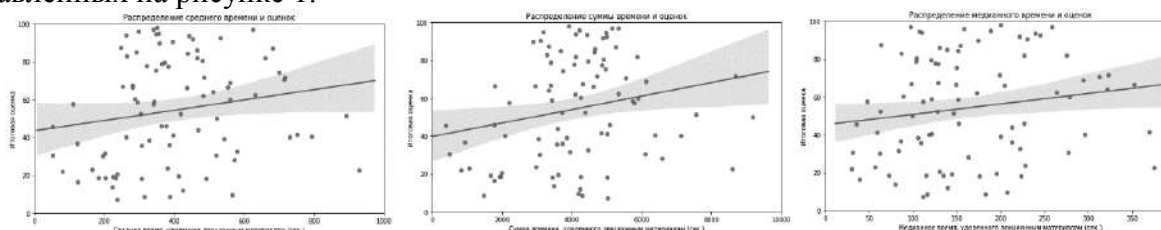


Рисунок 1. Графики рассеивания («Медианное время» и «Итоговая отметка», «Суммарное время» и «Итоговая отметка», «Медианное время» и «Итоговая отметка»)

Следующим этапом были оценены меры близости, с использованием в качестве объекта «Студента», в качестве первого признака - «Итоговая отметка», в качестве второго - «Среднее время», «Суммарное время» и «Медианное время», последовательно. Мера близости вычислялась с использованием расстояния Хэмминга. При использовании расстояния Хэмминга расстояния между объектами представляются в виде матрицы расстояний, которая представляет собой квадратную матрицу типа «объект - объект» (порядка $n - 4$), где в качестве элементов выступают расстояния между объектами в

метрическое объекте. Диагональные элементы такой матрицы равны нулю, так как расстояние между объектом и им самим равно нулю. В итоге были получены три графика, представленные ниже на рисунке 2, представляющие матрицы расстояний. Выявить каких-либо закономерностей не удалось.

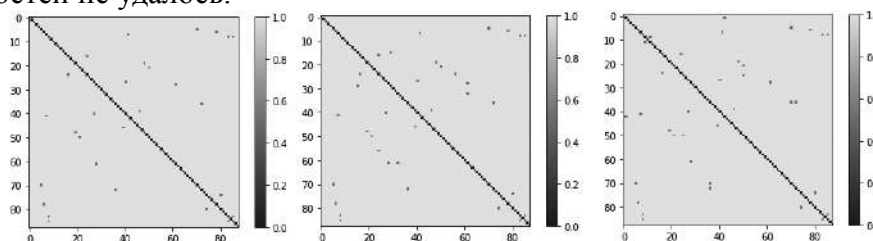


Рисунок 2. Графическое представление матриц расстояний («Среднее время - Итоговая отметка», «Суммарное время - Итоговая отметка», «Медианное время - Итоговая отметка»)

Заключение

В заключении авторы хотели бы обратить внимание на то, что в исследовании был использован открытый набор данных, в котором изначально присутствовали студенты, которые вообще не изучали теоретический материал или не сдавали итоговый экзамен. Полученные результаты могут быть интерпретированы как незаинтересованность студентов в изучении теоретического материала. Причиной этого могло быть короткое время сессий (всего полтора часа), за которые студент должен был выполнять ряд заданий (лабораторных работ) в среде обучения.

ЛИТЕРАТУРА

1. NMC Horizon Report > 2017 Higher education edition (Russian). Перевод: Московская школа управления СКОЛКОВО // NMC.org - New Media Consortium. 2017. URL:<https://www.nmc.org/publication/nmc-horizon-report-2017-higher-education-edition-russian/> (дата обращения: 21.11.2018)
2. CourseBurg. Как живут MOOC платформы в российских реалиях. Первое в России исследование платформ онлайн-образования // CourseBurg. 2017. URL: https://courseburg.ru/analytics/Issledovanie_MOOC_platform.pdf (дата обращения: 30.11.2018)
3. Меньшикова Е. «Как российские MOOC-платформы конкурируют с мировыми гигантами онлайн-образования, выяснили в Courseburg» // ITMO.News. 2018. URL: <http://news.ifmo.ru/ru/news/7692/> (дата обращения: 30.11.2018)
4. Patricia Witkin. 2016 Learning Trends on Udemy //Udemy Blog. 2016. URL: <https://about.udemy.com/instructors/2016-learning-trends-on-udemy/> (дата обращения: 27.11.2018)
5. Матросова Н.Д., Штенников Д.Г. Расширенная статистика LMS Moodle // Проблемы и перспективы внедрения свободного программного обеспечения в образовательных организациях Санкт-Петербурга. Мат-лы XI конференции. - 2018. - С. 16-21 (дата обращения: 22.11.2018)
6. Dua, D. and Karra Taniskidou, E. (2017). UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science. URL: <http://archive.ics.uci.edu/ml> (дата обращения: 13.11.2018)

7. M. Vahdat, L. Oneto, D. Anguita, M. Funk, M. Rauterberg.: A learning analytics approach to correlate the academic achievements of students with interaction data from an educational simulator // UCI Machine Learning Repository. 2015. URL: [https://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+\(EPM\)%3A+A+Learning+A+nalytics+Data+Set](https://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+(EPM)%3A+A+Learning+A+nalytics+Data+Set) (дата обращения: 13.11.2018)

Сведения об авторе

- | | | |
|--|---|---|
| Наталья Дмитриевна
Матросова | — | студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: aqice26@gmail.com |
| Дмитрий Геннадьевич
Штенников | — | доцент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: lisizina@corp.ifmo.ru |

Ссылка для цитирования: Матросова Н.Д., Штенников Д.Г. Исследование зависимости результатов тестирования от времени изучения теоретического материала студентами на открытом наборе данных // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 73—76.

**STUDY OF CORRELATION BETWEEN THEORY
LEARNING TIME AND TESTING RESULTS**

N.D. Matrosova, D.G. Shtennikov

ITMO University, 197101, St. Petersburg, Russia

Researchers tested the hypothesis of a correlation between online-course final grades and the learning materials studying time. The research was conducted using the open data set provided by the UCI Machine Learning Repository.

Keywords: MOOC, distance course, open data set, Hamming distance, Pearson correlation coefficient

Data on author

- | | | |
|------------------------------|---|---|
| Nataliia D. Matrosova | — | graduate student; ITMO University,
Faculty of Software Engineering and Computer Systems,
E-mail: ndmatrosova@corp.ifmo.ru |
| Dmitry G. Shtennikov | — | docent; ITMO University,
Faculty of Software Engineering and Computer Systems,
E-mail: dshtennikov@corp.ifmo.ru |

For citation: Matrosova N.D., Shtennikov D.G. Study of correlation between theory learning time and testing results // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 73—76 (in Russian).

УДК 004.89

ИЗВЛЕЧЕНИЕ КОНЦЕПТОВ ИЗ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ ТЕКСТОВ НА ОСНОВЕ КОНТЕКСТА

Л.А. Мирославская, И.А. Бессмертный

Университет ИТМО, 197101, Санкт-Петербург, Россия

В данной статье рассматривается возможность извлечения концептов из текстов, написанных на естественном языке. Мы описываем теоретическую часть комплексного подхода к извлечению концептов, основанную на векторном представлении контекста слова с целью повышения качества извлечения знаний из естественно-языковых текстов.

Ключевые слова: искусственный интеллект, обработка естественного языка, контекст слова, квантовый формализм, интеллектуальный анализ данных, извлечение информации, мешок слова, word2vec

Введение

С момента создания первого высокоуровневого языка программирования прошло более полувека. В этот период человек учился общаться с машиной на языке, понятном рукотворным вычислительным системам. В последнее время большую популярность набирает такой тренд, как «искусственный интеллект» (Artificial Intelligence), из которого сформировалось самостоятельное направление, получившее название «обработка естественного языка» (Natural Language Processing). Цель данного направления – научить машину понимать информацию, представленную в формате естественно-языковых (ЕЯ) текстов, а также извлекать смыслы и формировать новые знания. Основная задача обработки естественного языка – исследование методов, на основе которых будет возможным создание систем, реализующих ЕЯ общение между человеком и компьютером. Стоит отметить, что вопрос автоматической обработки текстов является междисциплинарным (в том числе, требует знаний лингвистики, математики, психологии) и является составляющей научной области известной, как компьютерная лингвистика (Computational Linguistics).

Обзор предметной области

К актуальным прикладным задачам компьютерной лингвистики относятся: машинный перевод, информационный поиск, аннотирование и классификация текстов, интеллектуальный анализ данных, извлечение информации.

В контексте данной работы, мы рассматриваем область извлечения информации (Information Extraction), как часть научного направления Data Mining. Обработка естественных языков – задача нетривиального характера. При ее решении возникают трудности, обусловленные языковой неоднозначностью, сложностью структуры, наличием ошибок в тексте, инверсивным порядком слов. Так, например, первая проблема выражается в явлениях полисемии, омонимии, синонимии. Полисемией называют наличие у одной единицы языка нескольких связанных между собой значений, в частности, полисемия слов, например: земля — суша, почва, конкретная планета [1]. Ключ к решению подобного затруднения – анализ контекста, в котором употребляется данное слово.

Стоит отметить, что на сегодняшний день, в обработке ЕЯ текстов не используются отдельные методы в чистом виде. Современный инструментарий представляет собой конгломерат разных подходов. Таким примером служит работа [2], в которой авторы предлагают для представления семантики текста использовать синтаксический и дистрибутивный подходы.

Наша задача заключается в определении концептов и подборе слов, посредством которых можно объяснить их смысловую нагрузку. Объектом данного исследования являются методы извлечения концептов из естественно-языковых текстов.

Предмет исследования – возможность использования квантового формализма для извлечения концептов из ЕЯ текстов.

В данной работе обсуждается теоретическая основа для определения концептов с учетом контекста и рассматривается решение следующих задач:

- Нормализация текста;
- Извлечения концептов посредством использования флективности языка с учетом контекста;
- Определение смыслов концептов на основе контекста.

Содержание исследования

Одним из методов представления слов через контекст является векторный подход, называемый мешком слов (bag-of-word). В данном случае после формирования тезауруса документа вектор составляется по следующему принципу: на позиции, соответствующей расположению слова в тезаурусе, выставляется значение, эквивалентное числу появлений термина в тексте. У данного подхода есть существенный недостаток – информация о взаимном расположении слов внутри текста отсутствует.

В 2013 году Томаш Миколов предложил свой подход, который получил название word2vec [3]. Его основа – гипотеза, которая в науке известна как гипотеза локальности – «слова, которые встречаются в одинаковых окружениях, имеют близкие значения». Суть метода заключается в том, что мы берем окно произвольного размера и через его призму рассматриваем отрывки текста. Размерность окна всегда должна быть нечетной. Это необходимо для того, чтобы слово, вектор которого мы строим, располагалось ровно в середине окна и имело одинаковое количество слов с обеих сторон. Такой подход дает нам возможность получить представление о том, какие слова встречаются чаще в одном контексте. Но мы все еще не знаем, как они связаны между собой.

Для более точного семантического моделирования текста используется алгоритм HAL (Hyperspace Analogue Language) [4]. В рамках такого подхода смысл кодируется счетчиками слов, которые входят в контекст, также учитывается расстояние между ними. Однако данный подход упускает порядок слов, что может негативно сказываться на качестве модели [5].

Для установления связи между словами мы будем использовать флективность языка [8]. Для этого нам необходимо провести лемматизацию кандидатов в термины. Таким образом, после проведения нормализации текста и применения HAL, мы получим представление документа в формате нормированной суммы векторов.

Тогда, используя квантовый формализм, мы можем сказать, что документ является состоянием квантовой частицы, представленной в виде суперпозиции возможных состояний - контекстов отдельных документов. Данный подход активно применяется во многих работах, посвященных анализу текстов с использованием формализма квантовой теории [5-7]. Так, например, в работе [5] для определения взаимосвязи между двумя словами, из которых состоит пользовательский запрос, в контексте одного документа используется алгоритм расчета параметра Белла.

Заключение

Результатом описанных выше манипуляций с текстом будет усредненный вектор, содержащий контекст – определение смысловой нагрузки выделенных концептов.

В данной статье мы описали теоретическую часть комплексного подхода к извлечению концептов, основанную на векторном представлении контекста слова. Далее необходимо разработать программное средство для моделирования экспериментов и провести анализ полученных результатов.

ЛИТЕРАТУРА

1. *Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С.* Автоматическая обработка текстов на естественном языке и анализ данных // Учебно-методическое пособие: 2017. 269 с.
2. *Sordoni A., Nie J., Bengio Y.* Modeling Term Dependencies with Quantum Language Models for IR // Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. — SIGIR '13. — New York, NY, USA : ACM, 2013 — P. 653–662
3. *Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean.* Efficient estimation of word representations in vector space. CoRR, abs/1301.3781
4. *Lund, K., Burgess, C.:* Producing high-dimensional semantic spaces from lexical co-occurrence. Behav Research Methods Instruments and Computers 28, pp. 203-208, (1996).
5. *Platonov A.V., Poleshchuk E.A., Bessmertny I.A., Gafurov N.R.* Using Quantum Mechanical Framework for Language Modeling and Information Retrieval // 12th IEEE International Conference on Application of Information and Communication Technologies, AICT 2018 - Conference Proceedings (Almaty, Kazakhstan 17-19 Oct 2018) - 2018, pp. 99-102.
6. *Wittek, P., Daranyi, S.:* Spectral composition of semantic spaces. QI 2011, LNCS 7052, Springer-Verlag: pp. 60-70, (2011).
7. *Barros J., Toffano Z., Meguebli Y., Doan B.* Contextual Query Using Bell Tests / — 2013 — 7 — Vol. Vol. 8369— P. 110–121.
8. *Bessmertny I.A., Platonov A.V., Poleshchuk E.A., Ma Pengyu* Syntactic Text Analysis Without a Dictionary // IEEE 10th International Conference on Application of Information and Communication Technologies (AICT) – 2016, pp. 1-3.

Сведения об авторе

**Люсьена Александровна
Мирославская**

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: lyusenamir25@gmail.com

**Игорь Александрович
Бессмертный**

— профессор; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: igor_bessmertny@hotmail.com

Ссылка для цитирования: *Мирославская Л.А., Бессмертный И.А.* Извлечение концептов из естественно-языковых текстов на основе контекста // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 77—80.

EXTRACTING CONCEPTS FROM NATURAL LANGUAGE TEXTS
BASED ON CONTEXT

L.A Miroslavskaya, I.A Bessmertny

ITMO University, 197101, St. Petersburg, Russia

In this article, the possibility of extracting concepts from texts written in natural language is considered. We describe the theoretical part of an integrated approach to the extraction of concepts, based on a vector representation of the context of a word with the aim of improving the quality of knowledge extraction from natural language texts.

Keywords: artificial intelligence, natural language processing, word context, quantum formalism, data mining, information extraction, bag-of-word, word2vec

Data on author

*Lyusiena A
Miroslavskaya*

— student; ITMO University;
Faculty of Software Engineering and Computer Systems,
e-mail: lyusenamir25@gmail.com

Igor A Bessmertny

— professor, ITMO University;
Faculty of Software Engineering and Computer Systems,
e-mail: igor_bessmertny@hotmail.com

For citation: *Miroslavskaya L.A. Bessmertny I.A.* Extracting concepts from natural language texts based on context // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 77—80 (in Russian).

УДК 004.934.2

ПОСТРОЕНИЕ АВТОМАТИЧЕСКИХ ДИАЛОГОВЫХ СИСТЕМ ДЛЯ СЛУЖБ ТЕХНИЧЕСКОЙ ПОДДЕРЖКИ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ БЕЗ УЧИТЕЛЯ

А.А. Нугманова^{1,3}, И.А. Черных^{1,2}, М.А. Ховричев^{1,3}, А.В.Булусева²

1 - Университет ИТМО, 197101, Санкт-Петербург, Россия

2 - ЦРТ-инновации, 196084, Санкт-Петербург, Россия

3 - ЦРТ, 196084, Санкт-Петербург, Россия

Работа выполнялась при финансовой поддержке Министерства
Образования и Науки Российской Федерации, Соглашение
14.575.21.0178 (RFMEFI57518X0178)

Разработаны и экспериментально исследованы методы построения автоматических диалоговых систем с применением методов машинного обучения без учителя. Преимуществом предложенных подходов является отсутствие необходимости использования ручной экспертной разметки обучающего корпуса диалоговых данных. Получены оценки качества результата диалоговой системы, построенной с использованием разработанных методов.

Ключевые слова: диалоговая система, обучение без учителя, системы на основе поиска

Введение

Построение автоматических диалоговых система для контактных центров и служб технической поддержки в настоящее время становится все более актуальной задачей в области обработки естественного языка [1].

Традиционный подход к построению автоматических диалоговых систем основан на методах обучения с учителем и базируется на трех основных шагах:

1. Построение диалогового сценария (дерева тематик или состояний);
2. Помещение в каждое состояние примеров пользовательских запросов;
3. Обучение классификатора.

В настоящий момент первые 2 шага предполагают большое количество ручной работы экспертов, которая не всегда возможна. Цель данной работы состоит в исследовании методов построения автоматических диалоговых систем на неразмеченных данных.

Обзор предметной области

Автоматические диалоговые системы можно разделить на 2 основных типа: диалоговые системы на основе поиска и генеративные диалоговые системы.

Основная идея методов диалоговых систем на основе поиска состоит в том, чтобы выбрать наиболее подходящий ответ путем ранжирования ответов-кандидатов. В качестве ответов-кандидатов обычно используются ответы из обучающей базы. Ранжирующим показателем при этом могут быть как сходства вопросов, так и линейные или нелинейные преобразования между векторными представлениями вопросов и ответов [2 - 4].

Генеративные диалоговые системы составляют ответ на каждый запрос в режиме реального времени. Ответ обычно генерируется пословно на основании скрытого вектора

вопроса [5]. Проблема таких подходов заключается в необходимости больших объемов данных для обучения.

Содержание исследования

Рассмотренная нами система работает на подходе на основе поиска, который предполагает выбор подходящего ответа из заранее подготовленной базы ответов. Архитектура системы была основана на сямской сети, которая так же известна под названием Двойной Кодировщик (Dual Encoder), представленная в работе [2]. Основная идея данной архитектуры заключается в том, что она учится предсказывать вероятность того насколько ответ подходит текущему контексту. Таким образом, сопоставляя текущий контекст диалога с каждым ответом из базы, система подбирает наиболее вероятный ответ. В контекст диалога входит вопрос пользователя и все предыдущие реплики.

Мы использовали архитектуру [2] в двух подходах и с двумя типами кодировщиков.

Первый подход включал оригинальную идею, описанную в [2, 3] и определял вероятность того насколько ответ подходит текущему контексту.

Во втором подходе архитектура с двумя кодировщиками использовалась только для извлечения признаков, описывающих вектор контекста. В данном подходе вместе с каждым ответом из базы хранится вектор контекста, в котором текущий ответ был использован. Наиболее вероятным считается тот ответ, соответствующий контекст которого ближе к контексту текущего вопроса. В качестве меры близости используется косинусное расстояние. В данном подходе мы так же рассмотрели два типа кодировщиков: основанный на рекуррентной сети с GRU ядром (DE GRU) и основанный на механизм внимания (DE ATT).

Все наши эксперименты проводились на русскоязычном корпусе, содержащем диалоги технической поддержки. Диалоги представляют собой человеческую беседу между пользователями и оператором. Корпус полностью неразмеченный и не содержит каких-либо ручных меток. Размер корпуса – 25000 диалогов со средней длиной 4 пары «вопрос-ответ».

Для оценки качества моделей была использована ручная разметка. Ручная разметка включала в себя базовый подход к оценке диалоговых систем на основе поиска – метрику $\text{recall}@k$, которая отражает долю правильных ответов, входящих в топ k ответов системы [2, 3, 4]. Для подсчета метрики система на каждый тестовый вопрос выдает k наиболее вероятных ответов и если подходящий ответ оказывается среди них, то данный вопрос помечается как правильный.

Тестовый набор данных был специально отобран экспертами и включал 400 вопросов. Вопросы содержали только целевые вопросы для данной системы, исключаящие приветствия, прощания и благодарности.

На каждый тестовый вопрос мы получали 3 наиболее вероятных ответа, после чего эксперты оценивали каждый из них. Для более точной оценки, каждому ответу они присваивали метку по четырехбалльной шкале. Каждую метку можно интерпретировать следующим образом: 0 – ответ совсем не подходит; 1 – ответ частично подходит; 2 – ответ содержит информацию, необходимую для правильного ответа; 3 – ответ является правильным.

На основании разметки, полученной от экспертов, подсчитывались 2 метрики:

CR (Correct Response) – $\text{recall}@3$ для правильных ответов. В этом случае вопрос помечается как правильный, если хотя бы на один из трех ответов есть оценка выше 1.

UR (Unsure Response) – $\text{recall}@3$ для неуверенных ответов. В этом случае вопрос помечается как правильный, если хотя бы на один из трех ответов есть оценка выше 0.

В Таблице 1 представлены результаты ручной разметки каждой из систем, основанных на 4х подходах. Последняя строка показывает результаты ансамбля из 3х систем (усреднение Word2Vec + Косинусное расстояние, признаки от DE RNN + Косинусное

расстояние, Признаки от DE АТТ + Косинусное расстояние). В режиме ансамбля вероятность каждого ответа представляла собой взвешенную сумму вероятностей от всех систем.

Таблица. Результаты систем с различными подходами.

Подход	R	R
Dual Encoder (DE)	0.12	0.45
Усреднение Word2Vec + Косинусное расстояние	0.28	0.74
Признаки от DE RNN + Косинусное расстояние	0.38	0.74
Признаки от DE АТТ + Косинусное расстояние	0.40	0.79
Ансамбль 3х систем	0.46	0.71

Заключение

В результате работы был проведен анализ подходов на основе поиска. Разработаны и экспериментально исследованы несколько методов построения автоматических диалоговых систем с применением методов машинного обучения без учителя. Основным преимуществом предложенных архитектур является отсутствие необходимости дополнительной ручной разметки данных для обучения системы.

На основании результатов из таблицы, лучшие результаты при оценке правильных ответов (CR) были достигнуты системой, основанной на ансамбле трех подходов, который для выбора ответа считает взвешенную сумму оценок от нескольких архитектур.

ЛИТЕРАТУРА

1. *Chakrabarti, Chayan, and George F. Luger.* "Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics." *Expert Systems with Applications* 42.20 (2015): 6878-6897.
2. *Lowe, Ryan, et al.* "The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems." *arXiv preprint arXiv:1506.08909* (2015).
3. *Kadlec, Rudolf, Martin Schmid, and Jan Kleindienst.* "Improved deep learning baselines for ubuntu corpus dialogs." *arXiv preprint arXiv:1510.03753* (2015).
4. *Lison, Pierre, and Serge Bibauw.* "Not All Dialogues are Created Equal: Instance Weighting for Neural Conversational Models." *arXiv preprint arXiv:1704.08966* (2017).
5. *Shang L., Lu Z., Li H.* Neural responding machine for short-text conversation // *arXiv preprint arXiv:1503.02364*. – 2015.

Сведения об авторе

Айгуль Айратовна Нугманова	—	аспирант; Университет ИТМО, Факультет информационных технологий и программирования; E-mail: nugmanova@speechpro.com
Ирина Александровна Черных	—	инженер; Университет ИТМО, Факультет информационных технологий и программирования; E-mail: chernykh-i@speechpro.com
Михаил Аркадьевич Ховричев	—	студент; Университет ИТМО, Факультет информационных технологий и программирования; E-mail: khovrichev@speechpro.com
Анна Владимировна Булушева	—	научный сотрудник, научно-исследовательский центр ЦРТ-инновации, E-mail: bulusheva@speechpro.com

Ссылка для цитирования: Нугманова А.А., Черных И.А., Ховричев М.А., Булушева А.В. Построение автоматических диалоговых систем для служб технической поддержки методами машинного обучения без учителя // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 81—84.

UNSUPERVISED TRAINING OF AUTOMATIC DIALOG SYSTEMS FOR THE CUSTOMER SUPPORT SERVICE

A.A.Nugmanova^{1,3}, I.A.Chernykh^{1,2}, M.A.Khovrichev^{1,3}, A.V.Bulusheva²

1 - ITMO University, 197101, St. Petersburg, Russia

2 - CRT-Innovation, 196084, St. Petersburg, Russia

3 - CRT, 196084, St. Petersburg, Russia

This work was financially supported by the Ministry of Education and Science of the Russian Federation, Contract 14.575.21.0178 (ID RFMEFI57518X0178)

Methods for constructing automatic dialog systems using machine learning methods without a teacher have been developed and experimentally investigated. The advantage of the proposed approaches is the absence of the need to use the expert hand labeling of the training data. The quality of developed dialogue system was estimated.

Keywords: dialog systems, unsupervised training, retrieval-based model

Data on author

Aigul A. Nugmanova	—	student; ITMO University, Faculty of Information Technologies and Programming; E-mail: nugmanova@speechpro.com
Irina A. Chernykh	—	engineer; ITMO University, Faculty of Information Technologies and Programming; E-mail: chernykh-i@speechpro.com
Mikhail A. Khovrichev	—	student; ITMO University, Faculty of Information Technologies and Programming; E-mail: khovrichev@speechpro.com
Anna V. Bulusheva	—	scientist; STC-Innovations; E-mail: bulusheva@speechpro.com

For citation: Nugmanova A.A., Chernykh I.A., Khovrichev M.A., Bulusheva A.V. Unsupervised training of automatic dialog systems for the customer support service // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 81—84 (in Russian).

УДК 004.07

УПРАВЛЕНИЕ ПРОСТРАНСТВЕННЫМИ РЕСУРСАМИ СИСТЕМ ХРАНЕНИЯ ДАННЫХ

Е.Д. Пойманова, Т.М. Татарникова

*Санкт-Петербургский государственный университет аэрокосмического приборостроения, 190000,
г. Санкт-Петербург, Россия*

Задача хранения данных не сводится только к выделению емкостей для размещения данных. Необходимы новые модели управления ресурсами систем хранения данных, учитывающие технологии многоуровневого хранения и проактивной миграции информации. Современные системы хранения данных по-прежнему остаются одномерными – многоуровневая миграция данных в них определяется только временем, прошедшим с момента последнего обращения к данным. В тоже время есть множество метрик, которые могут учитываться в моделях управления жизненным циклом информации, такие как скорость предоставления запрошенных данных, стоимость потери данных, период, через который информация переходит на следующую стадию. Соответственно, требуются алгоритмы и модели проактивной политики миграции данных с подтверждением результатов эффективного хранения данных согласно этой политике. В работе предлагается последовательность механизмов распределения данных по уровням системы хранения данных, горизонтального распределения на уровне RAID и миграции данных. Результат применения предлагаемых механизмов визуализируется в виде матрицы, размер которой соответствует количеству уровней хранения и количеству физических или логических носителей. Все механизмы управления ресурсами систем хранения данных основаны на анализе метаданных сохраняемых файлов.

Ключевые слова: многоуровневое хранение данных, система хранения данных, управление ресурсами систем хранения, миграция данных, жизненный цикл данных

Введение

Для хранения больших объемов данных используются системы хранения, представляющие собой совокупность оборудования и программного обеспечения, реализующих запись данных, физическое хранение и доступ к ним [1, 2].

Структура хранилища данных предполагает наличие трех уровней: RAID, автоматизированные библиотеки (лента, оптические диски), носители длительного хранения. Каждый уровень хранилища предполагает свои технологии хранения [3].

В свою очередь сохраняемые файлы характеризуются временем хранения, размером и частотой обращения к ним. При этом, частота обращения к файлам является одной из определяющих характеристик системы хранения данных (СХД): для данных, к которым обращение происходит часто необходимы носители с минимальным временем доступа, а для данных, к которым обращение происходит редко, например, к архивным данным такие требования к носителю не предъявляются [4].

Обзор предметной области

Распределение информации по уровням СХД должно происходить в зависимости от характеристик входящих данных и предполагает следующие механизмы:

- механизм выбора уровня хранилища в зависимости от времени хранения;
- механизм миграции данных по уровням в зависимости от частоты обращения к ним.

Кроме того, перед записью на определенный уровень хранилища предлагается проводить анализ данных с целью выбора оптимальной файловой системы (для уровня RAID) или типа архивного носителя, что позволит эффективно использовать пространство хранилища. Таким образом, хранилище данных будет иметь ячеистую структуру, где в каждой ячейке будут храниться данные с определенными характеристиками.

Важную роль в работе СХД играет файловая система. Файловую систему выбирают исходя из типа носителя, типа операционной системы, поддерживаемого размера логической единицы данных (кластера, блока), максимального размера файла, максимального размера тома, длины имени и других.

Содержание исследования

В работе предлагается механизм выбора уровня хранилища данных в зависимости от их типа. Осуществление такого распределения возможно на основании анализа организационных метаданных, содержащих сведения о типе данных.

Поскольку данные, сохраняемые в СХД разнородны, т.е. прежде всего имеют различный формат и, как следствие – размер, важно выбрать для них оптимальную файловую систему, рационально использующую пространство на носителях. В одной СХД может быть несколько файловых систем, расположенных в разных томах дискового хранилища или на разных логических уровнях системы [5].

На уровне RAID для эффективного использования пространственных ресурсов СХД предлагается проводить анализ сохраняемых файлов и, в зависимости от их размера, размещать их в разных томах RAID. Каждый том при этом должен иметь свою файловую систему с определенным размером логического блока данных. Соответственно, чем больше файл, тем больший логический блок для него является оптимальным.

На нижних уровнях хранилища предлагается разделение пространства по типам носителей, например, стример, DVD, BD для уровня автоматизированных библиотек и М-диск, стеклянный диск и ДНК – для носителей длительного хранения.

Частота обращения к данным является одной из характеристик, влияющих на распределение данных по ячейкам хранилища. После первоначального распределения файлов при записи в хранилище предлагается на основании статистических данных о частоте обращения к файлам λ осуществлять их миграцию между уровнями хранилища [6].

Предлагается последовательно осуществлять работу механизмов распределения данных по уровням, горизонтального распределения и миграции.

Первым этапом является распределение данных по уровням хранилища в зависимости от метаданных, содержащих тип сохраняемых файлов. Каждый уровень хранилища предполагает определенное время хранения данных и определенную частоту обращения. Чем ниже уровень, тем больше время хранения и меньше частота обращения к данным. Тип сохраняемых файлов носит исключительно субъективный характер и определяется при сохранении [7].

На втором этапе, после определения уровня хранилища, происходит горизонтальное распределение файлов по выбранному уровню в зависимости от размера файла с данными.

На третьем этапе происходит миграция данных по уровням хранилища в зависимости от частоты обращения к ним. Частота обращения определяется на основе статистики за определенный период. Могут учитываться пороговые значения частоты λ_1 и λ_2 , при преодолении которых данные мигрируют на соответствующий уровень хранилища.

Очевидно, что механизм миграции позволяет преодолеть недостатки субъективного выбора типа сохраняемых файлов, который был реализован на первом этапе.

Результатом предлагаемого механизма управления пространственными ресурсами в общем виде является матрица размера $m \times n$, где m – это количество уровней хранилища, а n – количество физических или логических носителей. Элементами матрицы являются множества файлов с данными, имеющие определенные значения характеристик: тип файла (type), размер файла (f). Частота обращения к данным (λ) при первоначальном размещении данных в хранилище не учитывается, поскольку в этот момент нет накопленной статистики обращения к данным.

Для распределения файлов по ячейкам матрицы предлагается проводить анализ метаданных входного потока файлов с использованием нейронных сетей Кохонена. Выбор данного метода анализа данных обусловлен особенностями алгоритма кластеризации в этих нейронных сетях: наличие неконтролируемого обучения, разбиение исходного множества объектов на классы в зависимости от значений признаков объектов, результат кластеризации визуализируется в виде топологической карты.

В случае предлагаемого распределения файлов по ячейкам хранилища данных очевидно, что выбранные характеристики файлов задают векторы значений признаков объектов. Таким образом, каждая ячейка получающейся карты должна соответствовать элементу матрицы управления пространственными ресурсами.

Заключение

Совокупность предложенных механизмов позволяет управлять емкостью хранилища и осуществлять рациональное использование носителей. Все механизмы основаны на анализе метаданных сохраняемых файлов.

ЛИТЕРАТУРА

1. *Farley M.* Building Storage Networks. — McGraw-Hall: Osborne, 2001.
2. *Landauer R.* Irreversibility and heat generation in the computing process // IBM Journal of Research and Development. 2000. Vol. 44. N 1. P. 261—269.
3. Information Storage and Management. 2nd Edition. — John Wiley & Sons Inc., 2016.
4. *Татарникова Т. М., Пойманова Е. Д.* Технологии долговременного хранения данных//Материалы Международной науч.-практической конференции «Наука и образование в XXI веке». Тамбов: Изд-во ТРОО «Бизнес-Наука-Общество». 2013. Ч. 31. С. 136—137.
5. *Kish L. B.* Moore's law and the energy requirement of computing versus performance // IEEE Proceedings: Circuits, Devices and Systems. 2004. Vol. 151, N 2. P. 190—194. *Татарникова Т.М.* Анализ данных. – СПб: СПбГЭУ, 2018.
6. *Kutuzov O. I., Tatarnikova T. M.* Model of a self-similar traffic generator and evaluation of buffer storage for classical and fractal queuing system. In Moscow Workshop on Electronic and Networking Technologies, MWENT. 2018. P. 1—3.
7. *Пойманова Е. Д.* Обобщенная модель процесса сохранения данных// Информационно-технологическое обеспечение цифровой экономики: сборник статей сотрудников кафедры информационных систем и технологий. — СПб.: Изд-во СПбГЭУ, 2018. С. 80—85.

Сведения об авторе**Екатерина Дмитриевна
Пойманова**— студент; Университет аэрокосмического приборостроения,
Кафедра безопасности информационных систем;
E-mail: e.d.poymanova@gmail.com**Татьяна Михайловна
Татарникова**— профессор; Университет аэрокосмического приборостроения,
Кафедра безопасности информационных систем;
E-mail: tm-tatarn@yandex.ru

Ссылка для цитирования: Пойманова Е.Д., Татарникова Т.М. Управление пространственными ресурсами систем хранения данных // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 85—88.

MANAGEMENT OF STORAGE SPACE RESOURCES***E.D. Poymanova, T.M. Tatarnikova****State University of Aerospace Instrumentation, 190000, St. Petersburg, Russia*

Storing data task is not limited only to allocation of buffers for data placement. New models of storage resource management are needed, taking into account multi-level storage and proactive migration of information. Modern data storage systems are still one-dimensional – multilevel data migration in them is determined only by time elapsed since the last access to data. At the same time, there are many metrics that can be taken into account in information life-cycle management models, such as the rate at which the requested data is provided, the cost of data loss, the period through which information passes to the next stage. Accordingly, algorithms and models of proactive data migration policy are required with confirmation of the results of effective data storage in accordance with this policy. The paper proposes a sequence of mechanisms for distributing data across storage systems, horizontal distribution at the RAID level, and data migration. The result of applying the proposed mechanisms is visualized in the form of a matrix, the size of which corresponds to the number of storage levels and the number of physical or logical media. All storage resource management mechanisms are based on the analysis of the metadata of the stored files.

Keywords: multi-level data storage, data storage system, storage resource management, data migration, data life cycle

Data on author***Ekaterina D. Poymanova***— student; State University of Aerospace Instrumentation,
Department of Security Information Systems;
E-mail: e.d.poymanova@gmail.com***Tatiana M. Tatarnikova***— professor; State University of Aerospace Instrumentation,
Department of Security Information Systems;
E-mail: tm-tatarn@yandex.ru

For citation: Poymanova E.D., Tatarnikova T.M. Management of storage space resources // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 85—88 (in Russian).

УДК 004.415.2

«PLR»: ПРОГРАММА ПРОВЕРКИ ЛАБОРАТОРНЫХ РАБОТ ПО КУРСУ «ОПЕРАЦИОННЫЕ СИСТЕМЫ»

В.Е. Шульмина

Университет ИТМО, 197101, Санкт-Петербург, Россия

С целью автоматизации процесса проверки правильности выполнения лабораторных работ разработана программа «PLR». Аналогичные продукты уже используются при обучении, однако все они имеют различия, зависящие от их предназначения. Данная программа написана на языке сценариев Bash и представляет собой набор модулей, каждый из которых отвечает за проверку конкретной лабораторной работы. Также в программе подразумевается наличие входных данных, принимаемых в качестве эталонных при проверке. Запуск программы осуществляется через терминал на каждой виртуальной машине, которую настраивал студент. Итоговыми результатами работы программы «PLR» являются сформированный текстовый документ, представляющий собой отчет, и сгенерированное значение контрольной суммы, подтверждающее подлинность предоставляемого отчета.

Ключевые слова: программа автоматизации, процесс контроля знаний, лабораторные работы, контрольная сумма, виртуальные машины, linux, язык программирования bash

Введение

В настоящее время стремительно развивается и модернизируется процесс обучения в высших учебных заведениях [1]. Используется множество инструментов для проверки и контроля знаний студентов. К таким инструментам относят электронные курсы, тесты и программы для проверки лабораторных работ. Результатом работы таких программ являются отчеты, на основании которых преподаватель принимает решение о зачете или незачете лабораторной работы, которую выполнял студент.

Использование подобных инструментов позволит ускорить процесс обучения, облегчив работу преподавателя [2]. Это также позволит сократить время проверки и снизить влияние человеческого фактора на процесс обучения [1]. За определенную единицу времени появляется возможность проверить большее число студентов, что является показателем более эффективного использования временных ресурсов. Однако сейчас в высших учебных заведениях на лабораторных и практических занятиях преподаватели продолжают индивидуальную проверку выполненных работ, что замедляет процесс проверки и может негативно сказываться на учебном процессе.

С целью автоматизации процесса проверки лабораторных работ и следующего за этим увеличения эффективности использования временных ресурсов (академических часов) была разработана и внедрена в учебный процесс программа «PLR».

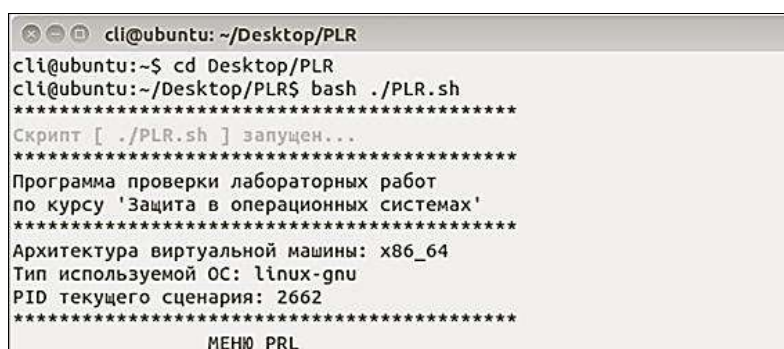
Обзор предметной области

Перед непосредственной разработкой алгоритма работы программы и написанием программного кода, были изучены существующие на данный момент аналоги. В ходе их анализа было выявлено, что между обозреваемыми аналогами существуют отличия, зависящие от направленности программ.

Одни программы производят проверку с набором тестов, другие проверяют соответствие с заданными условиями [3, 4]. Также программы пишутся под определенную операционную систему, что исключает возможность использования программы на других платформах [5].

Содержание исследования

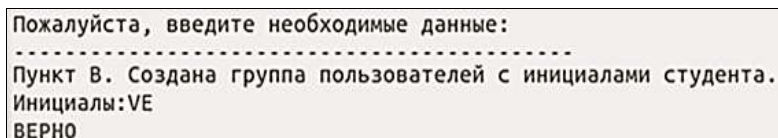
Для написания программы, которая будет применяться в операционной системе Ubuntu, был выбран язык сценариев Bash. Выбор языка обусловлен возможностью использования интерактивной командной оболочки, способной читать данные из терминала пользователя и в последующем писать данные в этот терминал [6]. Таким образом реализуется удобный интерфейс для взаимодействия пользователя с программой (рисунок 1). При выборе оболочки Bash также была учтена возможность последующей интеграции и использования программы в других операционных системах семейства Linux.



```
cli@ubuntu: ~/Desktop/PLR
cli@ubuntu:~$ cd Desktop/PLR
cli@ubuntu:~/Desktop/PLR$ bash ./PLR.sh
*****
Скрипт [ ./PLR.sh ] запущен...
*****
Программа проверки лабораторных работ
по курсу 'Защита в операционных системах'
*****
Архитектура виртуальной машины: x86_64
Тип используемой ОС: linux-gnu
PID текущего сценария: 2662
*****
МЕНЮ PRL
```

Рисунок 1. Интерфейс программы «PLR»

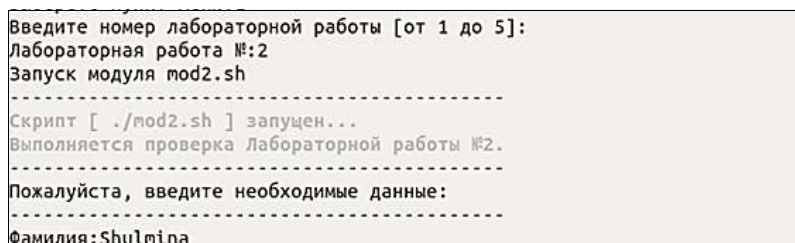
В программе подразумевается наличие входных данных, используемых для последующей проверки настроек виртуальных машин в зависимости от задания на лабораторную работу, в качестве эталонных (рисунок 2).



```
Пожалуйста, введите необходимые данные:
-----
Пункт В. Создана группа пользователей с инициалами студента.
Инициалы:VE
ВЕРНО
```

Рисунок 2. Пример использования входных данных при проверке

Программа представляет собой набор модулей. Каждый из них отвечает за проверку конкретной лабораторной работы. Модули связаны общей «оболочкой», в которой есть возможность выбора определенного модуля (рисунок 3). Запуск программы осуществляется через терминал на каждой виртуальной машине, настраиваемой студентом.



```
Введите номер лабораторной работы [от 1 до 5]:
Лабораторная работа №:2
Запуск модуля mod2.sh
-----
Скрипт [ ./mod2.sh ] запущен...
Выполняется проверка Лабораторной работы №2.
-----
Пожалуйста, введите необходимые данные:
-----
Фамилия:Shulmina
```

Рисунок 3. Выбор модуля проверки лабораторной работы

Начало работы с программой «PLR» выглядит следующим образом:

1. После выполнения лабораторной работы студенту следует перенести на рабочий стол виртуальной машины каталог «PLR», в котором содержится

основной сценарий программы, а также дополнительные модули, отвечающие за проверку лабораторных работ.

2. Далее необходимо открыть терминал и выполнить несколько команд. Первой командой осуществляется переход в каталог «PLR», второй командой – запуск основного сценария программы (рисунок 4).

```
cli@ubuntu: ~/Desktop/PLR
cli@ubuntu:~$ cd Desktop/PLR
cli@ubuntu:~/Desktop/PLR$ bash ./PLR.sh
*****
Скрипт [ ./PLR.sh ] запущен...
*****
```

Рисунок 4. Запуск сценария «PLR.sh»

3. После запуска основного сценария рекомендуется следовать инструкциям в окне терминала для успешного выполнения проверки лабораторных работ и корректного завершения работы программы.

Заключение

Одним из итоговых результатов выполнения программы проверки является сформированный текстовый документ, представляющий собой отчет, содержащий общую информацию о студенте (дата формирования отчета, фамилия студента) и подробную информацию о проверке лабораторной работы (задания лабораторной работы, команды для проверки пунктов, результаты выполнения команд) (рисунок 5).

```
lw1_Shulmina.txt x
Дата: Fri May 25 22:40:53 PDT 2018
Студент: Shulmina
-----
Пункт№1 Задать имя, состоящее из фамилии обучаемого.
Для проверки использовалась команда HOSTNAME.
Требуемое имя ОС: Shulmina
Настроенное имя: ubuntu
:Ошибка!
```

Рисунок 5. Пример отчета, формируемого программой «PLR»

Другим результатом выполнения программы проверки является сгенерированное значение контрольной суммы (рисунок 6). Полученное значение является важным аспектом проверки подлинности отчета и используется преподавателем для принятия решения о зачете или незачете лабораторной работы.

```
ВНИМАНИЕ! НЕОБХОДИМО СДЕЛАТЬ СКРИНШОТ КОНТРОЛЬНОЙ СУММЫ:
Check sum: 5b9bc218637d7af0a0b4f67b1c4126ad lw3_Shulmina.txt
Скрипт [ ./mod3.sh ] завершил работу...
```

Рисунок 6. Сгенерированное значение контрольной суммы

Внедрение программы в учебный процесс обладает рядом достоинств:

1. облегчение процесса контроля знаний студентов;
2. предоставление достоверной информации о проверке;
3. понятный интерфейс для работы пользователей.

Стоит отметить, что программа представляет собой продукт с открытым исходным кодом, что дает возможность последующей ее модернизации и расширения для интеграции на другие платформы с учетом их особенностей.

ЛИТЕРАТУРА

1. *Самойлова И. А.* Сложность реализации автоматизированной проверки дипломных работ на соответствие техническим нормам. Казань: ООО «Издательство Молодой ученый», 2016. – С. 149-152.
2. *Касаткина, Н. Э.* Современные образовательные технологии в учебном процессе вуза: методическое пособие / Н. Э. Касаткина, Т. К. Градусова, Т. А. Жукова, Е. А. Кагакина, О. М. Колупаева, Г. Г. Солодова, И. В. Тимонина. – Кемерово: ГОУ «КРИПО», 2011. – С. 237.
3. *Долгушин, Н. А.* Использование системы conteste для проведения олимпиад по программированию / Н. А. Долгушин, М. Н. Оленькова. Москва: ООО "Информационно-технический отдел Академии Естествознания", 2016. – С. 290.
4. *Легалов, А. И.* Формальные языки и трансляторы: учебное пособие / И. А. Легалов, Д. А. Швец. – Красноярск, 2007.
5. Система проведения турниров по программированию Ejudge. [Электронный ресурс]: – Режим доступа: <https://ejudge.ru>
6. *Гаррельс, М.* Руководство по Bash для начинающих. [Электронный ресурс]: – Режим доступа: <http://rus-linux.net>

Сведения об авторе

**Валерия Евгеньевна
Шульмина**

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: shulminamix@mail.ru

Ссылка для цитирования: Шульмина В.А. «PLR»: программа проверки лабораторных работ по курсу «операционные системы»// Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 89—93.

PLR: AUTOMATED ATTESTATION OF LABORATORY WORKS ON THE COURSE "OPERATING SYSTEMS"**V.E. Shulmina***ITMO University, 197101, St. Petersburg, Russia*

In purpose to automate the process of checking the correctness of laboratory work, the program "PLR" was develop. Similar products are using in education, but they all have differences depending on their purpose. This program was write in the Bash scripting language, which allows the use of an interactive command shell that allows you to read data from a user terminal and then write data to this terminal. A program includes several modules, where each module is responsible for checking specific laboratory work. These modules are connect by a common "shell". In this "shell", you can select the required module. The program is launch via the terminal on each virtual machine that the student has set up. The program also implies the presence of input data, which is true value during the check. The results of the "PLR" program are the generated text document and the generated checksum value, which confirms the authenticity of the report.

Keywords: program automation, control knowledge process, laboratory work, checksum, virtual machines, Linux, Bash

Data on author***Valeriya E. Shulmina***

— student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: shulminamix@mail.ru

For citation: *Shulmina V.E.* "PLR": automated attestation of laboratory works on the course "operating systems" // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 89—93 (in Russian).

КОМПЬЮТЕРНАЯ БЕЗОПАСНОСТЬ

УДК 681.518.5

УРОВНИ БЕЗОПАСНОСТИ IOT

И.М. Брысина, Т.А. Маркина

Университет ИТМО, 197101, Санкт-Петербург, Россия

Целью исследования является определение оптимального способа защиты для данных IOT устройств. Для этого была проанализирована архитектура IOT, определены уровни защиты информации. Согласно существующим стандартам обеспечения безопасности IOT предполагается на уровнях: приложения, сети и устройства. В ходе проведённого анализа было определено, что обеспечение информационной безопасности IOT на уровне сети является наиболее эффективным, были определены дальнейшие решения для обеспечения безопасности IOT.

Ключевые слова: IOT, информационная безопасность, сетевая безопасность, "умный дом", центры обработки данных

Введение

Целью исследования уровней безопасности IOT является определение оптимального способа защиты для данных устройств. Под IOT понимают предоставление широкого спектра услуг на основе взаимодействия устройств друг с другом с помощью телекоммуникационных средств. Задачей исследования является определение концепции IOT, понимание, на каком из уровней IOT защита информации будет наиболее оптимальной. Благодаря этому появится представление, как создать корректную и наиболее эффективную концепцию защиты информации. Устройства IOT широко распространены в данное время, по прогнозам количество устройств в дальнейшем будет только увеличиваться. Это приведёт к изменениям в социальных, экономических и других сферах жизни [1]. Таким образом, тема актуальна в настоящее время.

Обзор предметной области

В данный момент согласно стандартам, существует эталонная модель IOT, которая включает в себя следующие уровни [2]:

- уровень приложения;
- уровень поддержки услуг и поддержки приложений;
- уровень сети;
- уровень устройства.

При этом общие возможности обеспечения безопасности предполагаются на уровнях: приложения, сети и устройства.

Содержание исследования

На уровне устройства очень сложно комплексно обеспечить безопасность.

Существует ряд основных трудностей:

- Пользователи не всегда вовремя устанавливают обновления. Устройство может получать обновление: от разработчиков/провайдеров автоматически или при помощи ручной

установки обновления. При этом оповещения об имеющихся обновлениях пользователи чаще всего не получают [3], а разработчики не всегда предоставляют возможность автоматического обновления.

- Невозможно закрыть все найденные уязвимости с помощью обновлений. Для уменьшения стоимости устройств их чаще всего производят из известных составляющих, но используют недорогие компоненты не самой последней версии аппаратного и программного обеспечения. При этом чаще всего обновления для старых версий уже не выпускают.

- Единых требований и стандартов не было долгое время. Таким образом в данный момент на рынке имеется много устройств, не соответствующих требованиям безопасности. Производители сами решали, как будет функционировать их устройство, какими протоколами и стандартами пользоваться. Заменить те устройства, которые уже есть на рынке, на те, которые бы подходили под стандарты, не представляется возможным.

- Существует вероятность наличия недеklarированных возможностей устройства (например, это может быть скрытый микрофон или передача данных с видеокамеры) [4]. Наличие чётких стандартов не гарантирует безопасность.

Таким образом, для обеспечения безопасности на уровне устройств придётся проверять все продаваемые устройства, это будет приносить лишние расходы, а также не имеет смысла из-за большого количества ИОТ. Установленные в данный момент устройства невозможно проверить на соответствие.

При обеспечении безопасности на уровне приложений следует помнить, что практически любое приложение и платформа уязвима, так как при написании программы невозможно сделать это безошибочно. Существуют несколько известных платформ, предоставляющих свои услуги по объединению ИОТ в единую сеть на программном уровне. При этом не существует платформы, которая не находила бы уязвимости в программном обеспечении [5]. Так же является проблемой то, что известные платформы обслуживают большое количество конечных устройств, и найденная уязвимость может привести к массовым атакам. Многие разработчики пренебрегают обновлениями, возраст программного обеспечения также может не соответствовать возрасту устройства. При проведении исследований домашних маршрутизаторов, было установлено, что программное обеспечение на них в среднем старше самого устройства на 4-5 лет [6]. На данном уровне ответственность за программное обеспечение и его взаимодействие с устройством ложится на производителя. Невозможность контроля за обновлениями и устранениями уязвимостей всех производителей приложений является источником нескончаемых угроз.

При обеспечении безопасности на сетевом уровне у нас есть возможность контролировать входящий и исходящий поток информации, анализировать данные. В данном случае мы сможем выстроить политику, при которой будет запрещено всё, что не разрешено, производить анализ трафика. При обеспечении сетевой безопасности не имеют значения уязвимости устройства и приложения. Так как передача данных уже стандартизирована протоколами, привести её к общей концепции будет намного проще.

Таким образом наиболее важным и менее затратным является обеспечение безопасности именно на уровне сети.

По прогнозам Cisco всего за 6 лет количество подключенных ИОТ вырастет в 2 раза и к 2021 году составит 13,7 млрд [7]. Это способствует увеличению трафика, проходящего через существующие центры обработки данных. При этом в зависимости от объема информации растёт стоимость на обеспечение информационной безопасности [8]. Весь трафик проходит через ограниченное количество серверов, обработка данного трафика может быть неточной, также при удачной целевой атаке на ЦОД, может быть получен доступ ко всем устройствам. Также встаёт вопрос о глобальной перезагрузке сетей [9].

Для того, чтобы уменьшить трафик, проходящий через ЦОД необходимо создать локальные центры обработки данных. Они должны самостоятельно анализировать поступающую информацию, не должны быть связаны друг с другом (для уменьшения вероятности массовых атак). Увеличение количества центров позволит им быстро обрабатывать поступающую информацию. Данные центры могут располагаться в зоне жизнедеятельности пользователя и контролироваться пользователем (как "умный дом"), могут быть распределены географически, контролироваться операторами связи.

При этом необходимо применение на практике единой, стандартизированной концепции защиты информации [10].

Заключение

В ходе исследования было установлено, что обеспечение безопасности на сетевом уровне IOT будет наиболее эффективно. При этом необходимо создать локальные ЦОД, для уменьшения объёма трафика, уменьшения вероятности массовых атак. В дальнейших планах развития создание устройства, способного обеспечивать безопасность IOT на уровне LAN сети.

ЛИТЕРАТУРА

1. *Markeeva A.V.* Internet of things (IOT): opportunities and threats for contemporary organizations. [Электронный ресурс]: <https://cyberleninka.ru/article/v/internet-veschey-iot-vozmozhnosti-i-ugrozy-dlya-sovremennyh-organizatsiy>
2. ITU-T Y.4000/Y.2060 (06/2012)
3. *Krebs B.* IoT Reality: Smart Devices, Dumb Defaults [Электронный ресурс]: <https://krebsonsecurity.com/2016/02/iot-reality-smart-devices-dumb-defaults>
4. *Матюха Д. В.* Перехват управления устройством интернета вещей // Сборник трудов IX научно-практической конференции молодых ученых "Вычислительные системы и сети (Майоровские чтения)"
5. Amazon устранила уязвимости IoT, влияющие на их умные устройства [Электронный ресурс]: <https://www.make-info.com/amazon-patched-vulnerabilities-in-their-smart-devices/>
6. *Schneier B.* The Internet of Things Is Wildly Insecure—And Often Unpatchable . [Электронный ресурс]: https://www.schneier.com/essays/archives/2014/01/the_internet_of_thin.html
7. CISCO [Electronic resource]: https://www.cisco.com/c/ru_ru/about/press/press-releases/2018/02-08.html
8. *Евдокимов И.В., Алалван А.З.Д., Тимофеев Н.А., Нехоношин С.Р.* Интернет вещей в контексте экономики программной инженерии и управления стоимостью проекта // Интернет-журнал «Науковедение». Том 9, №6 (2017)
9. *Ромасевич Е.П.* Исследование агрегированного трафика беспроводных IOT устройств// Современные информационные технологии и ИТ-образование. Том 13, №3 (2017)
10. *Лисецкий Ю.М., Бобров С.И.* Новые угрозы информационной безопасности или оружие массового заражения // "Математичні машини і системи". №1 2018

Сведения об авторе

- Ирина Михайловна
Брысина** — студент; Университет ИТМО,
Факультет программной инженерии и компьютерных технологий;
E-mail: brysina.i.m@gmail.com
- Татьяна Анатольевна
Маркина** — старший преподаватель; Университет ИТМО,
Факультет программной инженерии и компьютерных технологий;
E-mail: markina_t@corp.ifmo.ru

Ссылка для цитирования: Брысина И.М., Маркина Т.А. Уровни безопасности ИОТ // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 94—97.

IOT SECURITY LEVELS**I.M. Brysina, T.A. Markina***ITMO University, 197101, St. Petersburg, Russia*

The purpose of the study is to determine the optimal method of data security IOT devices. For this, IOT architecture was analyzed, information security levels were determined. According to existing standards, IOT security assumed at levels: applications, networks, and devices. In the course of the analysis, it was determined that providing of IOT's information security at the network level is the most effective, further solutions were determined to ensure the security of IOT.

Keywords: IOT, information security, network security, smart home, data processing centers

Data on author

- Irina M. Brysina** — student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: brysina.i.m@gmail.com
- Tatyana A. Markina** — senior lecturer; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: markina_t@corp.ifmo.ru

For citation: Brysina I.M., Markina T.A. IOT security levels // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 94—97 (in Russian).

УДК 004.056

АНАЛИЗ МЕТОДОВ ЗАЩИТЫ ИНФОРМАЦИИ, ПЕРЕДАВАЕМОЙ ПО ВОЛОКОННО-ОПТИЧЕСКИМ ЛИНИЯМ СВЯЗИ

А. Кулмаханов

Университет ИТМО, 197101, Санкт-Петербург, Россия

В данной работе исследованы методы обеспечения защиты информации в волоконно-оптических линиях связи. Определены узкие места данных методов, а также перечислены способы борьбы с угрозами, которые могут возникнуть в канале связи. По результатам сравнительного анализа выявлено, что методы квантовой криптографии обеспечивают более высокую степень защиты от перехвата информации на линиях связи.

Ключевые слова: оптическое волокно, защита информации, помехозащищенность, перехват трафика, системы мониторинга ВОЛС, защищенная сеть передачи данных, квантовая криптография

Введение

Угрозы, основанные на оптическом контакте, можно разделить на пассивные и активные.

Пассивные угрозы направлены в основном на несанкционированное использование ВОЛС, не оказывая при этом влияние на ее функционирование [1]. Активные угрозы используются с целью нарушения нормального функционирования ВОЛС, путем целенаправленного воздействия на ее компоненты. К активным угрозам можно отнести, например, нарушение работы линии связи, утечка информации, НСД к информации и т.д.

В работе рассматриваются методы, предусматривающие защиту, как от пассивных атак, так и от активных.

Обзор предметной области

Считается, что ВОЛС, в силу особенностей распространения электромагнитной энергии в оптическом волокне (ОВ), обладают повышенной скрытностью, но на сегодняшний день имеется множество доказательств, что это не так [2]. Поэтому разработка методов защиты информации, передаваемой по ВОЛС, является актуальной задачей. Утечка информации, передаваемой через ВОЛС в большинстве случаев возможна в следствие перехвата трафика и речевой информации субъекта. В работе рассмотрены методы, которые позволяют избежать утечки информации.

Содержание исследования

Метод непрерывного мониторинга позволяет контролировать характеристики оптического волокна, точки соединения, места сопряжения различных кабелей, а также возможные неоднородности в кабелях и т. д. Данная задача контроля за сетью передачи данных решается с помощью автоматизированной системы оптических кабелей, которая представляет собой систему удаленного контроля оптических волокон (Remote Fiber Test System - RFTS). Удаленный контроль оптических волокон выполняется с помощью оптических импульсных рефлектометров (Optical Time Domain Reflectometer - OTDR), которые осуществляют диагностирование волокон. Архитектура RFTS содержит следующие компоненты:

- аппаратную часть;
- систему управления;
- геоинформационную систему (ГИС) привязки топологии сети к карте местности;
- базу данных, в которой содержится информация об оборудовании сети, оптических кабелей.

Оптический рефлектометр периодически снимает данные по затуханиям с подключаемых к нему оптических волокон сети, каждая полученная рефлектограмма сравнивается с эталонной, которая отражает обычно исходное состояние волокна. Если отклонения от нормы превышают определенные заранее установленные пороги, то соответствующий блок автоматически посылает на центральный сервер системы предупреждение и сообщение о неисправности. Таким образом, система мониторинга обеспечивает повышенную безопасность сети передачи данных, так как любое несанкционированное подключение к волокну неизбежно приводит к дополнительным потерям в оптическом канале, что значит, будет обнаружено и зафиксировано системой в реальном масштабе времени.

Использование данного метода существенно сокращает время поиска неисправностей и упрощает проведение профилактического обслуживания ВОЛС. Учитывая размеры современных сетей передачи данных, важность и объемы передаваемой информации по ним, данный метод обеспечивает безопасную передачу данных, защиту от несанкционированного доступа к сети [3].

Следующий метод заключается в использовании специальной структуры волоконно-оптического кабеля. Делается это с целью затруднения механической разделки кабеля и преграждения доступа к оптическому волокну. В данном методе защиты используется многослойное оптическое волокно со специальной структурой отражающих и защитных оболочек. Попытки проникнуть к сердцевине кабеля обнаруживаются по изменению уровня контрольного (шумового) сигнала. Место несанкционированного доступа определяется с высокой точностью с помощью рефлектометра.

Основным недостатком данного метода является отсутствие информации о координатах появляющихся неоднородностей, что не дает возможности проводить более тонкий анализ изменений режимов работы ВОЛС для линий связи, имеющих большую протяженность [3].

Метод шифрования подразумевает использование алгоритмов квантовой криптографии. Один из надежных способов сохранить в тайне телефонные переговоры или информацию, передаваемую по компьютерным сетям. Суть метода заключается в использовании квантов света (фотонов) для защиты информации [4]. Исходя из этого, наибольшее практическое применение квантовой криптографии находится в сфере защиты информации, передаваемой по волоконно-оптическим линиям связи (ВОЛС). Данный метод осуществляется за счет передачи фотонов на большие расстояния с минимальными искажениями. В качестве источников фотонов применяются лазерные диоды передающих модулей ВОЛС. Таким образом, образуется квантовый оптический канал (КОКС). В связи с тем, что КОКС не обладают такой скоростью как у ВОЛС, данный канал используется для осуществления криптографических методов, то есть он применяется для распределения ключей, которые затем используются средствами шифрования высокоскоростного потока данных [4]. При переходе от сигналов КОКС в сигналы ВОЛС и вступают в действие законы квантовой физики. Именно на использовании этих законов в сочетании с методами классической криптографии основывается квантовая криптография. При использовании данного метода попытка перехвата информации из квантового канала связи неизбежно приводит к внесению в него помех, обнаруживаемых санкционированными пользователями. Алгоритм данного метода состоит из стека протоколов, реализация которого позволяет

законным пользователям обеспечить формирование общего ключа, при условии утечки к злоумышленнику не более заданного количества информации или отказ от сеанса передачи при невыполнении этого условия [5].

Преимуществом данного метода является доступная реализация при больших длинах линии связи. Таким образом, создание защищенных корпоративных или локальных сетей различных топологий с использованием методов квантовой криптографии является оптимальным решением.

Заключение

В статье подробно рассмотрены современные методы защиты информации, передаваемой по ВОЛС, а также механизмы их действия. В связи с тем, что оптическое волокно является основным способом передачи информации в компьютерных сетях, оно подвергается атакам со стороны злоумышленников. Поэтому данная задача актуальна на сегодняшний день, из-за чего возникает потребность в защите.

ЛИТЕРАТУРА

1. Техничко-коммерческий сборник. М., АО ВОТ N1, 2003 // “Волоконно-оптическая техника”.
2. Ross F. E., Hamstra J. R. “Forging FDDI”, IEEE J. Select Areas.
3. Stallings W., 2009 Cryptography and Network, 13-19.
4. Кулин С. Я., Хорошко Д. Б., Низовцев А. П.: Квантовая криптография: идеи и практика, 2007 // Т.15. – 177-191с.
5. Кулин С. Я. Квантовая информация: //учеб. для вузов. Академия, 2007 – Т.29 – 507-527с.

Сведения об авторах

Азамат Кулмаханов

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерных технологий;
E-mail: azacool77@gmail.com

Ссылка для цитирования: Кулмаханов А. Анализ методов защиты информации, передаваемой по волоконно-оптическим линиям связи // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 98—101.

ANALYSIS OF METHODS FOR PROTECTING INFORMATION TRANSMITTED VIA FIBER-OPTIC COMMUNICATION LINES

A. Kulmakhanov

ITMO University, 197101, St. Petersburg, Russia

In this paper, researched the methods of providing the protection of information in fiber-optic communication lines. Identified bottlenecks of these methods, as well as ways to deal with threats that may arise in the communication channel. According to the results of a comparative analysis, it was revealed that quantum cryptography methods provide a higher degree of protection against information interception on communication lines.

Keywords: fiber optic communication line, quantum cryptography, protecting information, optical fiber, Petersburg national research university

*Data on authors**Azamat Kulmakhanov*

— student; ITMO University,
Faculty of Software Engineering and Computer Engineering;
E-mail: azacool77@gmail.com

For citation: *Kulmakhanov A.* Analysis of methods for protecting information transmitted via fiber-optic communication lines // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 98—101 (in Russian).

УДК 004.031.4

КОНЦЕПЦИЯ СЕТЕВОЙ БЕЗОПАСНОСТИ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМ ПРОЦЕССОМ (АСУ ТП)

М.А. Пакулин

Университет ИТМО, 197101, Санкт-Петербург, Россия

В статье рассматривается вопрос сетевой защиты автоматизированных систем управления процессами, а также выявляются основные угрозы и базовые модели злоумышленников. Используя полученные данные, разработана концепция безопасности и рассмотрены основные способы защиты.

Ключевые слова: автоматизированная система управления процессами, промышленный интернет вещей, информационная безопасность, сетевая безопасность, информационные технологии

Введение

На сегодняшний день все большую популярность приобретает Индустриальный Интернет вещей (Industrial Internet of Things - IIoT) - система объединенных компьютерных сетей и подключенных промышленных (производственных) объектов со встроенными датчиками и ПО для сбора и обмена данными, с возможностью удаленного контроля и управления в автоматизированном режиме, без участия человека. IIoT представлен в промышленной, энергетической и транспортной отраслях. Реализация индустриального интернета вещей в промышленности сегодня – интеграция автоматизированных систем управления технологическим процессом (АСУ ТП) с другими технологиями. Как и любое нововведение, вовлечение ранее автономных АСУ ТП в Промышленный интернет несет с собой новые угрозы. Для их нейтрализации требуется обнаружить основные угрозы и разработать концепцию безопасности.

Обзор предметной области

Для создания концепции защиты необходимо определить архитектурные особенности защищаемых АСУ ТП. Они сильно разнятся от вида производства, однако имеют и общие признаки: все они являются системами реального времени, в которых критичны задержки, потери данных, время реакции; недопустимы некорректные показатели, а также перезагрузки и отключения – все это несет за собой огромные финансовые и/или производственные потери и угрожает здоровью людей. Кроме того, базово любая АСУ ТП делится на три уровня: диспетчеризации, контроллеров, контрольно-измерительного оборудования.

Вовлеченные в Промышленный интернет АСУ ТП имеют повышенную уязвимость: к специализированным атакам, таким как черви Stux-net, Shamoon, взлом серверов и кража информации OaSyS SCADA Telvent (Schneider Electric) и другим инцидентам [1], на подготовку которых требуется несколько месяцев, высококлассных специалистов и знание особенностей конкретного производства, добавляется весь спектр уязвимостей для корпоративных и домашних систем (от вирусов-майнеров до Petya), т.к. большинство автоматизированных рабочих мест используют Windows, на котором установлено специализированное ПО.

В случае с сетевой безопасностью АСУ ТП подразумевается, что она должна обеспечивать не только защиту канала, но и защиту от угроз, выполняемых с

использованием сетевой инфраструктуры, т.е. любых атак, выполняемых аутсайдером и той части атак, что может быть выполнена с использованием локальной сети. К сожалению, к обычным видам злоумышленников, характерным для корпоративных систем, инсайдеру и аутсайдеру, в случае АСУ ТП добавляется вид «халатный работник». По сути, это инсайдер, не несущий злого умысла, но своими действиями нарушающий работу АСУ ТП.

Помимо особенностей объекта и видов злоумышленников в реальной жизни свои важные коррективы вносят нормативно-правовые документы. Помимо общемирового опыта в виде серии ISO/IEC 27000 «Information technology – Security techniques – Information security management systems» и серии стандартов ISA/IEC 62443 «Security for Industrial Automation and Control Systems», часть которых принята в России в качестве ГОСТ Р [2], готовятся поправки в приказ ФСТЭК №31[3], ужесточающие некоторые из его положений. Разрабатываемая концепция должна удовлетворять их требованиям.

Содержание исследования

Важно понимать, что в случае защиты АСУ ТП, речь идет не о защите информации как таковой, а о защите процесса, останов которого приводит к любого рода потерям. Опираясь на требования нормативно-правовых документов, имея в виду перечисленные модели злоумышленников и учитывая особенности архитектуры АСУ ТП, можно сделать вывод, что в случае защиты АСУ ТП главными свойствами защиты являются целостность и доступность, а не конфиденциальность. Любая задержка или неверные показания с датчиков могут привести к фатальным последствиям на линии производства, тогда как показания датчиков не принесут прямой выгоды злоумышленнику. Объектами защиты при этом являются три уровня АСУ ТП, а также взаимодействие между ними и доступ к каждому из них. Еще одним важным постулатом в обеспечении защиты АСУ ТП является пресечение (чем радикальнее, тем лучше) любых возможностей, лежащих за рамками обязанностей работника.

Популярными путями реализации полученной концепции являются контроль полосы пропускания (датчикам, как правило, не требуется широкий канал), контроль доступа в сети ПоТ [4], списки доступа, реализация верхнего уровня только на доверенных узлах, повсеместные файерволы на верхнем и среднем уровнях архитектуры АСУ ТП, максимальная сегментация сети, минимализация точек входа в сеть АСУ ТП, системы одностороннего межсетевого взаимодействия, системы обнаружения и предотвращения атак со специализированным для АСУ ТП набором сигнатур атак, файеролы прикладного уровня и т.д. [5, 6]

Построение системы с учетом накопленного опыта позволяет минимизировать отдельные недостатки компонентов системы защиты. Грамотная архитектура локальной сети не только полностью изолирует уровень контрольно-измерительного оборудования от уровня диспетчеризации, а значит и от возможных атак аутсайдера, но и в том числе сводит к минимуму возможности инсайдера, оставляя только одно АРМ доступным для атаки из локальной сети. Однако брандмауэры сводят к минимуму успех атаки по данному направлению.

Заключение

При данных условиях остается только одно слабое место в сетевой инфраструктуре АСУ ТП – соединение между нодой уровня диспетчеризации, доступной для инсайдера из той же подсети, и нодой уровня контроллеров, находящейся в другой подсети. В лабораторных условиях планируется воспроизведение данной ситуации с помощью сетевого датчика температуры, ПО для передачи данных и компьютера, имитирующего работу АРМ. В ходе работы определяются детали взаимодействия (порты, протоколы и пр.), после чего

будет произведена попытка нарушения работы. По итогам анализа инцидента будут предложены рабочие варианты защиты от подобных ситуаций.

ЛИТЕРАТУРА

1. *Грицай Г., Тиморин А., Гольцев Ю., Ильин Р., Гордейчик С., Карпин А.* Безопасность промышленных систем в цифрах v2.1*. [Электронный ресурс]: http://www.ptsecurity.ru/download/SCADA_analytics_russian.pdf Доступ свободный. Дата обращения 28.11.2018 г.
2. ГОСТ Р ИСО/МЭК 15408-1-2002. Информационная технология (ИТ). Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. [Текст]. – Москва: ИПК Изд-во стандартов, 2004. – 36 с.
3. Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды [Электронный ресурс]: Приказ Федеральной службы по техническому и экспортному контролю от 14 марта 2014 г. № 31 – Режим доступа: Система Гарант
4. *Аникеенко В.* Безопасность АСУ ТП и контроль привилегированных пользователей. [Электронный ресурс]: <http://www.anti-malware.ru/node/11899> Доступ свободный. Дата обращения 28.11.2018 г.
5. *Лукацкий А.* Обзор мировых стандартов ИБ АСУ ТП и советы по их применимости в российских условиях. [Электронный ресурс]: <http://www.slideshare.net/lukatsky/ss-14591209> Доступ свободный. Дата обращения 28.11.2018 г.
6. *Мальнев, А.* Защита АСУ ТП: от теории к практике [Текст] // "Information Security/ Информационная безопасность". - 2012. - № 6. – С. 24-26.

Сведения об авторе

Максим Андреевич Пакулин — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники ;
E-mail: pakulin.maksim@mail.ru

Ссылка для цитирования: *Пакулин М.А.* Концепция сетевой безопасности в автоматизированных системах управления технологическим процессом (АСУ ТП) // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 103—105.

THE CONCEPT OF NETWORK SECURITY IN AUTOMATED PROCESS CONTROL SYSTEMS

M.A. Pakulin

ITMO University, 197101, St. Petersburg, Russia

The paper discusses the issue of network protection of automated process control systems, as well as identifies the main threats and basic models of intruders. Using the data obtained, a security concept was created and the main ways of protection were considered.

Keywords: automated process control system, process control system, electronic resource, industrial internet, information technology, information security

*Data on author**Maksim A. Pakulin*

— student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: pakulin.maksim@mail.ru

For citation: *Pakulin M.A.* The concept of network security in automated process control systems // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 103—105 (in Russian).

УДК 004.49

МЕЖСЕТЕВЫЕ ЭКРАНЫ УРОВНЯ ВЕБ-ПРИЛОЖЕНИЯ В СОВРЕМЕННОМ МИРЕ

Т.С. Платонов, А.А. Оголюк

Университет ИТМО, 197101, Санкт-Петербург, Россия

Рассмотрены типы атак на веб-приложения и требования к современным межсетевым экранам уровня веб-приложения. Существующие решения были проверены на соответствие требованиям. Представлен проект межсетевого экрана уровня веб-приложения, связанный с информацией из профилей пользователей веб-приложения.

Ключевые слова: межсетевой экран уровня веб-приложения, отказ в обслуживании, взлом, дерево решений, цифровой отпечаток, интернет-мошенничество

Введение

Веб-приложения страдают от трёх различных типов атак. Самая разрушительная из них — это отказ в обслуживании или DDoS. Программные или аппаратные межсетевые экраны обычно расположены на четвёртом, транспортном уровне по модели OSI [1]. Они фильтруют IP- или MAC- адреса, обнаруживают нетипичный трафик и т.д. Основная проблема с такими межсетевыми экранами в том, что пользователь не может сообщить о ложном срабатывании, потому что браузер показывает ошибку сети, а не веб-страницу. В таком случае каждое ложное срабатывание дорого обходится для репутации веб-приложения. В дополнение к репетиционным рискам, хранение в быстрой, оперативной памяти чёрного списка с IPv6 адресами пока представляет нерешённую проблему.

Второй тип атак взламывает само веб-приложение, и не пытается сделать веб-приложение недоступным. Это может быть как атака типа SQL-инъекция, межсайтовый скриптинг, подбор пароля или другие атаки, направленные на хищение данных. Для предотвращения данных атак обычно используются межсетевые экраны уровня веб-приложений на основе правил [2, 3]. Но даже с обновлённой базой правил, она остаётся статичной, и новые методы атак обходят защиту [4].

Последний тип атак — это интернет-мошенничество, потому что главная ценность веб-приложений — это данные или пользователи. Вредоносные пользователи или боты могут вести себя почти как нормальные пользователи, но с некоторыми, специфичными для приложения шаблонами поведения. Это исследование направлено на смягчение атак данного типа.

Обзор предметной области

Современные межсетевые экраны уровня веб-приложения хорошо борются с атаками первых двух типов. Согласно обзору (<https://www.anti-malware.ru/compare/web-application-firewall-waf-2017>) за 2017 год от информационно-аналитического центра Anti-Malware.ru, на рынке межсетевых экранов уровня веб-приложения представлено достаточное количество продуктов, в том числе PT Application Firewall, Imperva Web Application Firewall, F5 Web Application Firewall. В каждом из этих продуктов представлено как средство по борьбе с атакой типа отказ в обслуживании, так и средства по борьбе со взломами, которые включают в себя известные правила для атак и механизмы создания новых правил для новых атак.

Существует и свободный продукт ModSecurity, который имеет большую базу правил и позволяет создавать собственные правила определения вредоносности запроса.

Однако, ни один из продуктов не интегрируется с самим веб-приложением, в результате межсетевые экраны вынуждены без знания конкретных алгоритмов работы веб-приложения предугадывать действия пользователей и определять их подозрительность. Такой подход не может быть эффективным, особенно с учётом изменений алгоритмов работы веб-приложений.

Содержание исследования

Люди совершают действия в веб-приложении с заметной задержкой в отличие от компьютеров. В дополнение к этому, действия пользователей обычно имеют некоторый порядок. Рассмотрим веб-приложение “галерея фотографий” с функцией поиска по похожим фотографиям. Сперва пользователь входит в веб-приложение под своей учётной записью, затем открывает список фотографий, а после этого открывает фотографию из списка и выполняет поиск фотографий, похожих на данную.

Боты обычно запрограммированы на конкретную цель с конкретным планом выполнения, поэтому они могут войти в веб-приложение под своей учётной записью, затем открыть список фотографий, после извлечь список интересующих его идентификаторов фотографий, а после выполнить поиск фотографий, похожих на фотографии по их идентификаторам. В таком случае шаг “открыть фотографию из списка” пропущен, и временной интервал между действиями будет необычно низким.

Ориентированный граф со средним интервалом на дугах может быть построен на основе потока взаимодействия пользователя с веб-приложением. Пользователи, которые не подходят под граф для конкретной страны/города/интернет-провайдера рассматриваются как подозрительные и они направляются на испытание в виде капчи, доказательства выполнения работы или предъявление второго фактора двухфакторной аутентификации.

Описанный выше метод подходит для веб-приложений без разделения на фронтенд- и бекенд- части. В случае, когда веб-приложение — это API, поток взаимодействия пользователя для различных клиентских приложений может сильно отличаться и не иметь системности, метод перестанет работать как предполагалось. Для решения этой проблемы, межсетевому экрану уровня веб-приложения следует собирать гораздо больше данных о профиле пользователя [5]. Эта информация включает в себя: данные о браузере (user-agent, доступные шрифты, заголовки запросов по-умолчанию, алгоритмы шифрования, используемые в TLS, язык), данные об операционной системе и аппаратном обеспечении (разрешение и диагональ экрана, версия операционной системы, видеокарта, MAC-адрес), данные, относящиеся к запросу (IP-адрес, интернет-провайдер, страна/город, RTT запроса), данные о пользователе (домен электронной почты, оператор мобильной связи, профили в социальных сетях и других веб-приложениях, типичное время активности, другая информация, относящаяся к доменной модели веб-приложения) и т.д. [6]. Набор этих данных также известен как цифровой отпечаток пользователя, и он практически уникален для каждого пользователя веб-приложения.

После сбора данных, каждый запрос или профиль в текущем состоянии обрабатывается с использованием дерева решений для определения, является ли он подозрительным или нет.

Заключение

Предложенный способ обнаружения подозрительных запросов в совокупности с тесной интеграцией с веб-приложением позволяет лучше справляться с интернет-мошенничеством. Известно два пути развития концепции. Во-первых, для обработки запросов в реальном времени деревья решений для профилей пользователей должны строиться в фоновом

режиме. Во-вторых, для улучшения качества классификации можно использовать деревья решений с градиентным бустингом.

Конечная цель исследования — это создание межсетевого экрана уровня веб-приложения, который может работать с использованием перечисленных методов с разными веб-приложениями. Он будет принимать общую информацию о профиле пользователя и действия пользователя. Программа будет рассчитывать репутацию профиля пользователя, и на основе неё определит, является ли профиль подозрительным.

ЛИТЕРАТУРА

1. *Fatkieva R.R.* Development of metrics for attack detection on the basis of network traffic analysis // Bulletin of the Buryat State University. Mathematics, Informatics. 2013. Vol. 1. P. 81-86.
2. *Endraca A., King B., Nodalo G., Maria M. Sta., Sabas I.* Web Application Firewall (WAF) // International Journal of e-Education, e-Business, e-Management and e-Learning. 2013. Vol. 3, No. 6. P. 451-455.
3. *Baranov P.A., Beybutov E.R.* Securing information resources using web application firewalls // Business Informatics. 2015. Vol. 4(34). P. 71-78.
4. *Melnikov V.G., Trifanov A.V.* WAF Bypass // Interexpo GEO-Siberia. 2017. Vol. 1. P. 113-117.
5. *Gómez-Boix A., Laperdrix P., Baudry B.* Hiding in the Crowd: an Analysis of the Effectiveness of Browser Fingerprinting at Large Scale // 27th International World Wide Web Conference. 2018. Vol. 1. P. 1-10.
6. *Vastel A., Laperdrix P., Rudametkin W., Rouvoy R.* FP-STALKER: Tracking Browser Fingerprint Evolutions // 39th IEEE Symposium on Security and Privacy. 2018. Vol. 1. P. 728-741.

Сведения об авторе

**Тимур Сергеевич
Платонов**

— студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: tpltzn@gmail.com

**Александр Александрович
Оголюк**

— доцент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: xms2007@yandex.ru

Ссылка для цитирования: Платонов Т.С., Оголюк А.А. Межсетевые экраны уровня веб-приложения в современном мире // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 106—109.

WEB APPLICATION FIREWALLS IN THE MODERN WORLD

T.S. Platonov, A.A. Ogolyuk

ITMO University, 197101, St. Petersburg, Russia

Explored web-application attack types and requirements for the modern web application firewalls. Existing solutions were tested for compliance with the requirements. The draft of web application firewall bound to application-specific profile data was introduced.

Keywords: web application firewall, denial of service, breach, decision tree, fingerprinting, internet fraud

*Data on author**Timur S. Platonov*

— student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: tpltnz@gmail.com

Alexander A. Ogolyuk

— docent; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: xms2007@yandex.ru

For citation: *Platonov T.S., Ogolyuk A.A.* Web application firewalls in the modern world // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 106—109 (in Russian).

УДК 004.49

ОБЗОР ХАРАКТЕРНЫХ ШАБЛОНОВ РАБОТЫ ШИФРОВАЛЬЩИКОВ И СПОСОБОВ ИХ ОБНАРУЖЕНИЯ

Т.С. Платонов

Университет ИТМО, 197101, Санкт-Петербург, Россия

При помощи компьютерного эксперимента установлено, какими характерными чертами обладают вирусы-шифровальщики для операционных систем Microsoft Windows. Рассмотрено, как эти черты изменялись в исторической перспективе. Предложен способ обнаружения деятельности шифровальщика в системе.

Ключевые слова: шифровальщик, алгоритм шифрования, симметричное шифрование, асимметричное шифрование, файловая активность, сетевая активность

Введение

Согласно исследованию Лаборатории Касперского, количество пользователей, атакованных шифровальщиками в 2016 году, было больше в девять раз по сравнению с 2015 годом. Это почти полмиллиона человек. Антивирусные компании только начинают добавлять модули обнаружения шифровальщиков в свои продукты. Кроме того, эффективность этих модулей не всегда высока. В связи с этим, разработка метода для обнаружения активности шифровальщика в системе важна для всех пользователей компьютеров. В задачу исследования входит определение основных признаков работы шифровальщиков, и создание метода для обнаружения как уже известных, так и неизвестных вирусов в операционной системе.

Обзор предметной области

В связи с тем, что первое обнаружение шифровальщиков было в 2007 году, а массовые случаи заражения начались в 2013, существует не так много исследований на эту тему. В статье [1] автор рассказывает о шифровальщиках, которые существуют на тот момент и способах предотвращения заражения. Однако, спустя всего год произошли серьёзные изменения на рынке шифровальщиков: целью атаки теперь стали крупные компании, появился тип распространения “шифровальщик как услуга” [2].

Антивирусные компании пытаются противодействовать новым угрозам, но, согласно исследованиям [3-5], они не всегда преуспевают в борьбе. Пользователи плохо защищены от атак. Даже крупные игроки на рынке с хорошими финансовыми возможностями не могут побороть шифровальщиков.

В исследованиях [6, 7] авторы делают вывод, что наиболее эффективный способ обнаруживать неизвестные угрозы — это выявить поведенческие шаблоны в потоке данных, что решается машинным обучением. Это исследование подтвердило, что шифровальщики могут быть эффективно обнаружены с использованием методов машинного обучения.

Содержание исследования

Шифровальщики прошли длинный путь развития. В момент их первого появления в 2007 году они использовали слабые алгоритмы шифрования, которые на мощностях антивирусных компаний удаётся в разумное время дешифровать. Но чаще всего серьёзный криптоанализ не требовался, потому что ключ можно было извлечь или из самого

исполняемого файла шифровальщика, или из системы. Шифровальщики были скорее экспериментом в мире киберпреступности, нежели разработкой с конкретной целью. Немного позже преступники начали использовать сильную криптографию, которую даже антивирусные компании не могли дешифровать в разумное время.

Со временем стало понятно, что эта концепция одного ключа на исполняемый файл не подходит для заражения большого числа пользователей. Ключи быстро достаются антивирусным компаниям, а они распространяют ключи среди пострадавших. Для решения этой проблемы шифровальщики стали запрашивать ключи шифрования у командных серверов, теперь ключи уникальные для каждой жертвы.

Область шифровальщиков-вымогателей замечает серьёзный преступный мир, более квалифицированные преступники. Они разрабатывают эффективные и криптостойкие способы шифрования. Во-первых, используется два алгоритма шифрования, быстрый и малонадёжный, и медленный и надёжный. Шифрование - медленный и затратный процесс, особенно если есть цель остаться незамеченным в системе. Злоумышленники стали шифровать только потенциально ценные файлы - документы, фотографии, файлы 3D-моделей, файлы исходного кода, или же шифровали только первые несколько байт файла, в результате чего файл невозможно использовать. Пользователи поняли, что необходимо делать резервные копии, преступники ответили - теперь вирусы стирают точки восстановления системы и данные сторонних программ резервного копирования.

Преступный мир понял, что гораздо больше денег можно получить от компаний, а не от домашних пользователей. Участились атаки на компании с низкой IT-культурой: медицинские, нефтегазовые, транспортные, государственные. Другие группы создают плацдарм для преступников, которые хотят войти в область, предоставляя шифровальщики как услугу, в которой входящим необходимо только распространить вирус.

На основе этих сведений можно выделить следующие характерные черты шифровальщиков:

- сетевой запрос для получения ключей шифрования;
- высокая активность в файловой системе, особенно перезапись первых нескольких байт файлов;
- отключение встроенных механизмов безопасности: отключение служб теневого копирования и восстановления системы.

В результате, если неизвестный код пытается выполнить указанные операции, можно с большой долей вероятности утверждать, что код является шифровальщиком.

Заключение

Рассмотрены все стадии развития шифровальщиков за последние 12 лет. Выделены поведенческие шаблоны, позволяющие определить наличие шифровальщика в операционной системе. Антивирусные компании внедрили модули для противодействия шифровальщикам, сфера в целом научилась предотвращать массовое заражение. Шифровальщики достигли определённых высот, но пока не видно перспектив дальнейшего роста. Они будут занимать некоторую долю рынка, потому что это действенный способ заработка.

ЛИТЕРАТУРА

1. Drobotun E.B. The analysis of activity and development trends of malicious programs "file locker-encoder" // Programmnye produkty i sistemy (software & systems). 2016. №2 (114). P. 77-82.
2. Kaspersky Security Bulletin 2016. The ransomware revolution. Story of the year // https://securelist.com/files/2016/12/KSB2016_Story_of_the_Year_ENG.pdf

3. Rubenking N.J. The Best Ransomware Protection of 2017 // <https://www.pcmag.com/roundup/353231/the-best-ransomware-protection>
4. MRG Effitas. In-the-wild Ransomware Protection Comparative Analysis 2016 Q3 // https://www.mrg-effitas.com/wp-content/uploads/2016/07/Zemana_ransomware_detection.pdf
5. AVLab. Protection Test Against Ransomware Threats // https://avlab.pl/sites/default/files/68files/ENG_2016_ransomware.pdf
6. Babenko L.K., Tumoyan E.P., Tsyganok K.V., Anikeev M.V. Classification of malicious software based on behavior features // Izvestiya SFedU. Engineering sciences. 2012. №4. P. 50-59
7. Zhernakov S.V, Gavrilov G.N Detection of malicious software using classical and neural network classification methods // Proceedings of the Voronezh State University of Engineering Technologies. 2015. №4 (66). P. 85-92

Сведения об авторе

Тимур Сергеевич Платонов — студент; Университет ИТМО,
Факультет программной инженерии и компьютерной техники;
E-mail: tpltzn@gmail.com

Ссылка для цитирования: Платонов Т.С. Обзор характерных шаблонов работы шифровальщиков и способов их обнаружения // Сборник трудов X международной научно-практической конференции молодых ученых «Программная инженерия и компьютерная техника (Майоровские чтения)». 2019. С. 110—112.

OVERVIEW OF ENCODERS TYPICAL PATTERNS AND METHODS OF THEIR DETECTION

T.S. Platonov

ITMO University, 197101, St. Petersburg, Russia

With the help of computer experiment it is established, what behavioural patterns do encoders for Microsoft Windows operating systems have. Considered how these methods have been changing since encoders first appearance. A method for detecting the activity of an encoders in the operating systems is suggested.

Keywords: encoder, encryption algorithm, symmetric encryption, asymmetric encryption, file system activity, network activity

Data on author

Timur S. Platonov — student; ITMO University,
Faculty of Software Engineering and Computer Systems;
E-mail: tpltzn@gmail.com

For citation: Platonov T.S. Overview of encoders typical patterns and methods of their detection // Proceedings of the Majorov International Conference on Software Engineering and Computer Systems. 2019. P. 110—112 (in Russian).

УНИВЕРСИТЕТ ИТМО

Сборник трудов
X международной научно-практической конференции
«Программная инженерия и компьютерная техника
(Майоровские чтения)»

В авторской редакции
Редакционно-издательский отдел
Университета ИТМО
Зав. РИО
Подписано к печати
Заказ №
Тираж 50 экз.

Н.Ф. Гусарова

ISBN 978-5-7577-0602-3



9 785757 706023

Редакционно-издательский отдел
Университета ИТМО

197101, Санкт-Петербург, Кронверкский пр., 49