

**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Уфимский государственный нефтяной технический университет»**

при поддержке:
Российской академии естественных наук
Академии наук Республики Башкортостан
Общественной организации
«Профессионалы дистанционного обучения»
Ассоциации образовательных программ
«Электронное образование Республики Башкортостан»
Российского союза научных и инженерных
общественных объединений
Партнерского центра
международного сертификационного холдинга IMQ

Информационные технологии Проблемы и решения

У ф а
УНПЦ "Издательство УГНТУ"
2 0 2 2

Информационные технологии. Проблемы и решения. – Уфа: УНПЦ "Издательство УГНТУ", 2022. 3(20). 100 с.

Information technology. – Ufa: UNPC "USPTU Publishers", 2022. 3(20). 100 p.

Учредитель:

**ФГБОУ ВО Уфимский государственный
нефтяной технический университет**

2022, 3(20)

Издается с 2014 г.

РЕДКОЛЛЕГИЯ

Главный редактор

Р.Н. Бахтизин, первый проректор Уфимского государственного нефтяного технического университета, д-р физ.-мат. наук, профессор

Члены редколлегии

Ю.Н. Белоножкин, канд. экон. наук, доцент кафедры финансы и кредит Сочинского государственного университета

Й. Дарадке, доцент, заместитель декана факультета вычислительной техники и сетей Университета принца Саттама бин Абдулазиза (PSAU) - Королевство Саудовская Аравия (KSA)

Ф.У. Еникеев, д-р техн. наук, профессор кафедры вычислительной техники и инженерной кибернетики Уфимского государственного нефтяного технического университета

В.В. Ерофеев, д-р техн. наук, профессор, руководитель Челябинского регионального отделения РАЕН

Н.В. Корнеев, д-р техн. наук, профессор кафедры управления безопасностью сложных систем Губкинского университета, член-корр. РАЕН

И.М. Михайловская, ст. преподаватель кафедры вычислительной техники и инженерной кибернетики Уфимского государственного нефтяного технического университета

Е.А. Султанова, канд. техн. наук, доцент кафедры вычислительной техники и инженерной кибернетики Уфимского государственного нефтяного технического университета, член-корр. РАЕН

В.Н. Филиппов, канд. техн. наук, доцент кафедры вычислительной техники и инженерной кибернетики Уфимского государственного нефтяного технического университета, действительный член РАЕН

© ФГБОУ ВО «Уфимский государственный нефтяной технический университет», 2022

© Коллектив авторов, 2022

Полнотекстовая версия выпуска размещена в Научной электронной библиотеке elibrary.ru по ссылке:

https://elibrary.ru/title_about.asp?id=61250

Подробности на сайте: <http://vtik.net>

Отпечатано с готового электронного файла.

Подписано в печать 04.07.2022. Формат 60x80/16. Гарнитура «Таймс». Усл. печ. л. 5,81. Тираж 800 экз. Заказ 150.

Учебный научно-производственный центр "Издательство Уфимского государственного нефтяного технического университета"

Адрес учебного научно-производственного центра "Издательство Уфимского государственного нефтяного

технического университета": 450064, Российская Федерация, Республика Башкортостан, г. Уфа, ул. Космонавтов, 1.

Founder:

**FSBEU NE Ufa State Petroleum
Technological University**

2022, 3(20)

Published since 2014

EDITORIAL BOARD

Editor-in-Chief

R.N. Bakhtizin, Dr. of Physical and Mathematical Sci., Professor, First Vice-Rector of Ufa State Petroleum Technological University

Editorial Board Members:

Yu. N. Belonozhkin, PhD Economic Sci. Department of Finance and Credit Sochi State university

Dr. Yousef Daradkeh, Associate Professor and Assistant Dean for Administrative Affairs, Department of Computer Engineering and Networks, Prince Sattam bin Abdulaziz University (PSAU) - Kingdom of Saudi Arabia (KSA)

F.U. Enikeev, Dr. of Technical Sci., Professor of Department of Computer Science and Engineering Cybernetics Ufa State Petroleum Technological University

V.V. Yerofeyev, Dr. Sci. Professor, Head of the Chelyabinsk regional branch of RANS

N.V. Korneev, Dr. Tech. Sci., Professor of the Department of Safety Management of Complex Systems, Gubkin University, Corresponding Member of the Russian Academy of Natural Sciences.

I.M. Mikhaylovskaya, Senior Lecturer of Department of Computer Engineering and Engineering Cybernetics Ufa State Petroleum Technological University

E.A. Sultanova, PhD, Deputy Head of Department of Computer science and Engineering cybernetics Ufa State Petroleum Technological University, corresponding member RANS

V.N. Filippov, PhD, Deputy Head of Department of Computer science and Engineering cybernetics of Ufa State Petroleum Technological University, Full member of the RANS

ОГЛАВЛЕНИЕ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В НАУКЕ, ОБРАЗОВАНИИ И ПРОИЗВОДСТВЕ

Хлыбов А.В., Гиниятуллин В.М., Ермолаев Е.В. МЕХАНИЗМ УВЕЛИЧЕНИЯ РАЗМЕРА И КОЛИЧЕСТВА ЯДЕР В СВЕРТОЧНОМ СЛОЕ НЕЙРОННОЙ СЕТИ.....	5
Бакулина В.Э., Воронина А.В., Чернова Е.В. ОБЗОР СТАНДАРТОВ В ОБЛАСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ.....	10
Бурдуковский С.О. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ ОБЛАСТИ ГЛАЗ НА ФОТОГРАФИИ ЛИЦА ЧЕЛОВЕКА.....	15
Вердиева Н.Н. ИССЛЕДОВАНИЕ КАЧЕСТВА ДАННЫХ В ГРАЖДАНСКОЙ НАУКЕ С ПОМОЩЬЮ РЕШЕНИЙ INDUSTRY 4.0.....	20
Хайдошкина Ю.В., Наташкина Е.А. АЛГОРИТМЫ ПРАКТИЧЕСКОЙ ПРИВЯЗКИ БПЛА К МЕСТНОСТИ.....	25
Гавриленко С.К., Минасов Ш.М., Мифтахов Б.А. МОБИЛЬНОЕ ПРИЛОЖЕНИЕ СИСТЕМЫ АВТОМАТИЗИРОВАННОГО УЧЕТА ПОСЕЩЕНИЯ ЗАНЯТИЙ...	31
Каданцев М.Н., Султанова Е.А, Филиппова Е.В. ИСПОЛЬЗОВАНИЕ ЯЗЫКА ЗАПРОСОВ SQL ПРИ ПРОЕКТИРОВАНИИ ИНФОРМАЦИОННЫХ СИСТЕМ	38
Кирюшин О.В., Мухамедьяров А.А. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПЕЧИ В СРЕДЕ SIMINTECH.....	43

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ, УПРАВЛЕНИИ И БИЗНЕСЕ

Макарова Е.А. ПОДДЕРЖКА ПРОЦЕССОВ ИНФОРМАЦИОННОГО УПРАВЛЕНИЯ С ПОМОЩЬЮ ПРОГРАММНОГО СЕРВИСА ДЛЯ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ СЛАБОСТРУКТУРИРОВАННЫХ ТЕКСТОВЫХ ДАННЫХ НА РУССКОМ ЯЗЫКЕ	50
Муталлапов Р.Н., Шишкина О.Ю. ПРОГРАММНЫЙ МОДУЛЬ РЕАЛИЗАЦИИ ШАБЛОНОВ ДЛЯ СОЗДАНИЯ ДОКУМЕНТОВ В СИСТЕМЕ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА.....	57
Мовсесян У.А., Ткаченко А.Л., Федорова В.А. ПРИМЕНЕНИЕ СИСТЕМЫ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА В ЮРИДИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ ОРГАНИЗАЦИИ.....	64

МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ

Лапин А.Н., Минасов Ш.М., Широкова А.А. МОДЕЛИ И АЛГОРИТМЫ СБОРА ИНФОРМАЦИИ ДЛЯ ФОРМИРОВАНИЯ БАЗЫ ДАННЫХ СИСТЕМЫ МОНИТОРИНГА ЛЕСНЫХ ПОЖАРОВ.....	70
Ткаченко К.С. ПРОГРАММНОЕ УПРАВЛЕНИЕ КОМПЬЮТЕРНЫМИ УЗЛАМИ В УСЛОВИЯХ ИЗМЕНЯЮЩИХСЯ ПОТОКОВ ЗАЯВОК С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА НАЗИНА-ПОЗНЯКА.....	76
Овсянникова А.С., Белозеров А.Е. ДОБАВЛЕНИЕ ОБЪЕКТА В ГЕОСОЦИАЛЬНОЙ СЕТИ «MYTRIP».....	81

СЕТИ И ТЕЛЕКОММУНИКАЦИИ

Гаврилов И.С., Тутубалин П.И. РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ АНАЛИЗА РАЗМЕЩЕНИЯ АППАРАТНО-ПРОГРАММНЫХ КОМПОНЕНТОВ В АРХИТЕКТУРЕ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ..... 87

СИСТЕМЫ И МЕТОДЫ ЗАЩИТЫ ИНФОРМАЦИИ

Корнеев Н.В. ИМПОРТОЗАМЕЩЕНИЕ И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ ОБЪЕКТОВ ТОПЛИВНО-ЭНЕРГЕТИЧЕСКОГО КОМПЛЕКСА СЕГОДНЯ И КОМПЛЕКСНАЯ БЕЗОПАСНОСТЬ В БУДУЩЕМ..... 95

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В НАУКЕ, ОБРАЗОВАНИИ И ПРОИЗВОДСТВЕ

УДК 004.932

МЕХАНИЗМ УВЕЛИЧЕНИЯ РАЗМЕРА И КОЛИЧЕСТВА ЯДЕР В СВЕРТОЧНОМ СЛОЕ НЕЙРОННОЙ СЕТИ

THE MECHANISM OF INCREASING THE SIZE AND NUMBER OF KERNELS IN THE CONVOLUTIONAL LAYER OF A NEURAL NETWORK

Хлыбов А.В., Гиниятуллин В.М., Ермолаев Е.В.,
Уфимский государственный нефтяной технический университет,
ул. Космонавтов, 1, г. Уфа, Республика Башкортостан, 450064, Россия

A.V. Khlybov, V.M. Giniyatullin, E.V. Ermolaev,
Ufa State Petroleum Technological University,
Kosmonavtov Str., 1, Ufa, Republic of Bashkortostan, 450064, Russia

e-mail: brinkinvision@gmail.com

Аннотация. В настоящее время развитие нейронных сетей идет по пути увеличения сложности архитектур моделей. Для того чтобы скорость вычислений была приемлемой применяются маленькие ядра. Однако их количество может достигать нескольких сотен в одном слое. При интерпретации результатов таких сетей нет четкого понимания, в каком слое и коэффициенты какого элемента необходимо изменить. В работе предлагается способ подбора дискретных ядер сверточного слоя и дальнейшего увеличения их размера. Преимуществом большой размерности ядер является возможность не только захватить большое количество признаков исходного изображения (рецептивное пятно), но и сократить выходной вектор сверточного слоя. Использование таких ядер позволяет уменьшить необходимое количество сверточных ядер в слое. В результате исследований оказалось, что при увеличении ядер необходимо учитывать и шаг прохода ядра свертки по изображению. На маленьком шаге точность сети практически всегда увеличивается, на большом шаге – точность уменьшается. Очевидно, что одного ядра недостаточно для получения требуемой точности. Поэтому были проведены эксперименты по добавлению дополнительных ядер. Было выявлено, что точность растет с увеличением количества ядер.

Abstract. Currently, the development of neural networks is on the path of increasing the complexity of model architectures. In order for the calculation speed to be acceptable, small kernels are used. However, their number can reach several hundred in one layer. When interpreting the results of such networks, there is no clear understanding in which layer and which element coefficients need to be changed. The article suggests a method for searching discrete kernels of the convolutional layer and further increasing their size. The advantage of the large dimension of the kernels is the ability not only to capture a large number of features of the original image (receptive field), but also to reduce the output vector of the convolutional layer. The use of such kernels makes it possible to reduce the required number of convolutional kernels in a layer. As a result of the research, it turned out that when increasing the kernels, it is necessary to take into account the stride of the convolution kernel passing through the image.

At a small stride, the accuracy of the network almost always increases, at a large stride, the accuracy decreases in proportion to the stride size. Obviously, one kernel is not enough to obtain the required accuracy. Therefore, experiments were carried out to add additional kernels. It was found that the accuracy increases with the increase in the number of kernels.

Ключевые слова: сверточное ядро, шаг, размер, точность, рецептивное поле.

Keywords: convolutional kernel, stride, size, accuracy, receptive field.

Существует огромное количество самых разных архитектур сверточных нейронных сетей, начиная с самых маленьких (LeNet-5) и до современных сложных моделей (Inception-V4) [1-3]. Всех их объединяет использование в сверточных слоях от нескольких десятков до нескольких сотен ядер. Следствием огромного количества ядер является появление большого количества гиперпараметров, которые необходимо учитывать, и весовых коэффициентов, которые необходимо четко интерпретировать, чтобы понять причины возникающих ошибок.

В связи с ростом сложности структур моделей пропорционально растут и требования к вычислительным ресурсам. Поскольку вычислительные мощности не являются неограниченными, логично появление различных методик упрощения применительно к существующим нейронным сетям. В последнее время идет развитие методики прунинга (прореживания) нейронных сетей [4], которая заключается в том, что элементы (веса, нейроны), практически не влияющие на конечный результат, удаляются из обученной модели. Однако применение данной методики не распространено ввиду отсутствия четкой формализации самого процесса прореживания. Нет четких указаний, удаление какого элемента в модели возможно и не приведет к существенному падению точности.

Становится актуальной задача совершенствования, модернизации существующих искусственных нейронных сетей с целью уменьшения потребления вычислительных мощностей.

Целью работы является разработка механизма увеличения размера ядра и количества ядер в сверточном слое, его экспериментальная проверка.

Было выдвинуто предположение замены обучения сверточных слоев на процесс перебора/подбора работоспособного набора сверточных ядер с заранее заданными дискретными значениями под конкретную прикладную задачу.

Исследование проводилось на наборе рукописных цифр MNIST. Использовалась следующая структура нейронной сети: один сверточный слой с одним трехзначным ядром (-1, 0, +1) и выходной полносвязный слой, состоящий из 10 нейронов.

При проходе сверточного ядра, значения которого представляют собой числа из некоторого заранее ограниченного целочисленного диапазона (в нашем случае это -1, 0 и +1), по изображению на выходе получается двумерная матрица, но меньшей размерности, чем исходное изображение (операция свертки).

Основой алгоритма перебора является троично-сбалансированная система счисления. Перебор ядра размером 3×3 с тремя возможными значениями сводится к $3^9 = 19683$ вариантам. Ячейки матрицы сверточного ядра нумеруются, начиная с нуля, младший разряд слева вверху (рисунок 1). Таким образом, из каждого ядра свертки можно сформировать девятиразрядное троично-сбалансированное число (ТСС) [5]. На рисунке красным цветом выделено знакоместо троично-сбалансированного числа, черным цветом его троичное значение. Номер ядра – это десятичное представление полученного числа.

Ядро №0		
0^0	0^1	0^2
0^3	0^4	0^5
0^6	0^7	0^8

Рисунок 1. Пример ядра в виде троично-сбалансированного числа

Для дальнейшего проведения исследования в качестве базовой точности (ассигасы) выбрана точность по обучающей выборке. Осуществив процедуру полного перебора максимальная точность составила 92.49 %, что соответствует ядру №-5468 (рисунок 2).

Ядро №-5468		
-1^1	0^2	1^3
1^4	1^5	1^6
1^7	1^8	1^9

Рисунок 2. Ядро с максимальной точностью

Увеличение размера ядра позволяет сократить выходной вектор сверточного слоя и увеличить рецептивное поле восприятия. Для динамических случаев (видеопоток данных) необходимы уже не плоские ядра, а объемные. Полный перебор таких ядер практически невозможен, поэтому предлагается подбирать ядро малой размерности (2*2 или 3*3) с высокой точностью распознавания и затем увеличивать его размер.

Процедура масштабирования сводится к последовательному добавлению строк/столбцов к исходному ядру. Таким образом, получаются не оптимальные варианты ядер, их можно назвать работоспособными. Именованное получаемых таким образом ядер состоит в следующем: сначала дается номер исходного ядра, затем через точку с запятой номера перебираемых строк/столбцов.

В процессе масштабирования возможно менять длину шага прохода сверточного ядра по изображению. Ниже представлены результаты для двух случаев:

- 1) когда шаг остается постоянным и равен шагу для малого ядра;
- 2) когда шаг меняется пропорционально увеличению размера ядра (таблицы 1 и 2).

Таким образом, в случае с пропорциональным увеличением шага получаемое путем масштабирования ядро 6*6 именуется следующим образом: -5468; 4; 37; 39; 94; -41; 283 (рисунок 3). Процедура масштабирования с постоянным малым шагом приводит к другому ядру: -5468; -9; 28; -30; -35; -18; 208 (рисунок 4).

Таблица 1 – Масштабирование с пропорциональным увеличением шага

Размерность	Десятичное представление строки/столбца в виде ТСС	Шаг	Точность, %
4*3	4	(3, 2)	91,86
4*4	37	(3, 3)	91,59
5*4	39	(4, 3)	89,60
5*5	94	(4, 4)	88,57
6*5	-41	(5, 4)	87,11
6*6	283	(5, 5)	85,53

Таблица 2 – Масштабирование с постоянным малым шагом

Размерность	Десятичное представление строки/столбца в виде ТСС	Шаг	Точность, %
4*3	-9	(2, 2)	92,78
4*4	28	(2, 2)	92,82
5*4	-30	(2, 2)	92,89
5*5	-35	(2, 2)	92,80
6*5	-18	(2, 2)	92,94
6*6	208	(2, 2)	93,17

Ядро №-5468; 4; 37; 39; 94; -41; 283					
-1	0	1	37 1	94 1	283 1
1	1	1	1	0	0
1 -5468	1	1	0	1	1
4 0	1	1	1	1	1
39 1	1	1	0	1	1
-41 -1	1	1	1	1	1

Рисунок 3. Ядро, полученное путем масштабирования с пропорциональным увеличением шага

По результатам таблиц 1 и 2 видно, что при малом шаге смещения ядра точность сети увеличивается, при большом – наблюдается существенное падение точности. Таким образом, возникает необходимость в подборе такого шага и размера ядра, когда и точность является приемлемой для решения задачи, и при этом реализуются преимущества больших ядер (сокращение размерности выходного вектора сверточного слоя, увеличение рецептивного поля).

Ядро №-5468; -9; 28; -30; -35; -18; 208					
-1	0	1	28 1	-35 0	208 1
1	1	1	0	-1	0
1 -5468	1	1	0	-1	-1
-9 -1	0	0	1	0	-1
-30 -1	0	-1	0	1	0
-18 0	-1	1	0	0	1

Рисунок 4. Ядро, полученное путем масштабирования с постоянным малым шагом

Чтобы увеличить точность модели к исходному ядру 3*3, полученному в процессе перебора, можно добавить дополнительные ядра. В таблице 3 представлены номера

дополнительных ядер и прирост точности, на рисунке 5 представлены сами ядра. По отдельности эти ядра дают следующие значения точности: для ядра №2903 точность равна 90.60%, ядра №-8969 89.67 %, ядра №-3873 88.04 %.

Таблица 3 – Изменение точности от количества ядер

Кол-во ядер	Номер ядра	Точность, %
1	-5468	92,49
2	2903	96,16
3	-8969	97,58
4	-3873	98,31

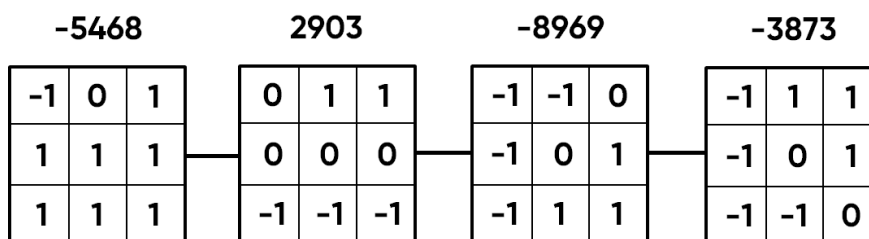


Рисунок 5. Дополнительные ядра 3*3

Очевидно, что, увеличивая количество ядер, можно добиться требуемой точности.

Выводы

Использование в качестве коэффициентов ядер сверточных слоев подбираемых, а не обучаемых ядер возможно и приводит к результатам с приемлемой точностью. Подбор ядра и использование дискретных коэффициентов позволит радикально упростить архитектуру модели и снизить требования к вычислительным ресурсам.

На длительность перебора влияют три составляющие: размер ядра, шаг свертки и количество ядер. Формулировка критерия их выбора позволит прогнозировать длительность перебора и объем необходимых ресурсов.

Литература

1. Alake R. Understanding and Implementing LeNet-5 CNN Architecture (Deep Learning) [Электронный ресурс]. 2020. URL: <https://towardsdatascience.com/understanding-and-implementing-lenet-5-cnn-architecture-deep-learning-a2d531ebc342> (дата обращения: 22.03.2022).
2. Inception-V4 and Inception-ResNets [Электронный ресурс]. 2022. URL: <https://www.geeksforgeeks.org/inception-v4-and-inception-resnets/> (дата обращения: 22.03.2022).
3. Васильев А.С., Батищева В.А., Филиппов В.Н., Некрасов Д.С. Актуализация решения проблем размещения зон пожарных депо в крупных городах с помощью технологии моделирования Электронный научный журнал Нефтегазовое дело. 2021. № 6. С. 136-145.
4. Tanaka H., Kunin D., Yamins D., Ganguli S. Pruning neural networks without any data by iteratively conserving synaptic flow [Электронный ресурс]. 2020. URL: <https://arxiv.org/abs/2006.05467> (дата обращения: 22.03.2022).

5. Buntime A. The Balanced Ternary Machines of Soviet Russia [Электронный ресурс]. 2016. URL: <https://dev.to/buntine/the-balanced-ternary-machines-of-soviet-russia> (дата обращения: 22.03.2022).

УДК 004.01

ОБЗОР СТАНДАРТОВ В ОБЛАСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

OVERVIEW OF SOFTWARE DEVELOPMENT STANDARDS

Бакулина В.Э., Воронина А.В., Чернова Е.В.,
ФГБОУ ВО «Магнитогорский государственный технический университет
им. Г.И. Носова»,
г. Магнитогорск, Российская Федерация

V.E. Bakulina, A.V. Voronina, E.V. Chernova,
FSBEI “Nosov Magnitogorsk State Technical University”,
Magnitogorsk, Russian Federation

e-mail: viktoriya.bakulina@mail.ru

Аннотация. В мире информационных технологий важно сохранить связь между различными областями взаимодействия: производителями, потребителями и экспертами по квалификации IT-продуктов. Научная статья посвящена теме стандартизации в области разработки программного обеспечения. Очень важно изучать стандарты, чтобы не произошло разрыва между ними, необходимо создать эффективный механизм взаимодействия всех сторон, создать необходимый уровень безопасности, обеспечиваемый продуктами ИТ. Программное обеспечение, представляет собой совокупность программ, которые обеспечивают функционирование, поддерживают стабильную работу компьютеров, а также позволяют с их помощью находить решение для задач предметных областей. Программное обеспечение составляет важную часть в компьютерной системе, является логическим продолжением технических средств и определяет сферу применения компьютера. Разработка же программного обеспечения представляет собой часть программной инженерии, которая отвечает за функционирование и сопровождение программных продуктов. В этой статье мы рассмотрели вопрос о существующих стандартах в области разработки программного обеспечения, кратко пояснили, что они из себя представляют и что включают. Также более подробно разобрали, в чем была необходимость развития стандартизации и создания стандартов для ИТ-продуктов. Были подведены итоги деятельности по изучению стандартизации.

Abstract. In the world of information technology, it is important to maintain the connection between different areas of interaction: manufacturers, consumers and experts in the qualification of IT products. The scientific article is devoted to the topic of standardization in the field of software development. It is very important to study standards so that there is no gap between them, it is necessary to create an effective mechanism for interaction between all parties, to create the necessary level of security provided by IT products. The software is a set

of programs that ensure the functioning, maintain the stable operation of computers, and also allow them to find solutions for problems of subject areas. Software is an important part of a computer system, is a logical continuation of technical means and determines the scope of the computer. Software development is a part of software engineering, which is responsible for the functioning and maintenance of software products. In this article, we examined the issue of existing standards in the field of software development, briefly explained what they are and what they include. They also explained in more detail what was the need to develop standardization and create standards for IT products. The results of standardization activities were summed up.

Ключевые слова: стандарты, стандартизация, программное обеспечение, ИТ-продукты, информационная безопасность.

Keywords: standards, standardization, software, IT products, information security.

Стандарты имеют большое значение – они предоставляют возможность разработчикам программного обеспечения изменять и применять данные и программы других разработчиков, осуществлять экспорт/импорт данных.

Главными итогами деятельности по стандартизации является увеличение степени соответствия продукта (услуги), процессов по их функциональному назначению, предотвращение технических барьеров в международном товарообмене, помощь научно-техническому прогрессу и совместной работе в различных сферах.

Все компании-разработчики должны гарантировать соответствующий уровень качества выпускаемого программного обеспечения (ПО). Для этих целей предусмотрены стандарты качества программного обеспечения или отдельные разделы в стандартах разработки программного обеспечения, посвященные требованиям к качеству программного обеспечения.

Такие стандарты обеспечивают взаимодействие между различными программами. Без таких стандартов программные продукты были бы «закрытыми» друг для друга.

Стандарты для написания документации делятся на два типа: Международные стандарты (ISO, IEEE Std); Советские и Российские ГОСТы (в обозначении Российских стандартов – символ Р).

Стандарты ISO 9000

Наиболее известной организацией, занимающейся проблемами стандартизации, является ISO (International Organization for Standardization). Россию в качестве члена ISO, представляет Госстандарт РФ (ГОСТ Р).

ISO представляет собой набор стандартов, которые помогают организациям обеспечивать удовлетворение потребностей клиентов и других заинтересованных сторон в рамках законодательных и нормативных требований, связанных с продуктом или услугой. ISO 9000 касается основ СМК, включая семь принципов управления качеством, лежащих в основе семейства стандартов.

Сфера деятельности ISO (ИСО) касается стандартизации во всех областях, кроме электротехники и электроники, относящихся к компетенции Международной электротехнической комиссии (МЭК).

Модели менеджмента качества ISO серии 9000 наиболее распространены во всем мире. Они универсальны, их можно применять независимо от того, в какой отрасли функционирует компания.

Основным сертифицируемым стандартом серии ISO 9000 является стандарт ISO 9001:2008, который представляет собой набор требований к системе менеджмента качества, используемых во всех организациях вне зависимости от вида деятельности, размера, принадлежности к государственному сектору, географического расположения.

Стандарт ISO 9000 определяет основные положения систем менеджмента качества, а также восемь принципов менеджмента качества, которые образуют основу для стандартов на системы менеджмента качества, входящих в серию стандартов ISO 9000. Стандарт ISO 9001 содержит набор требований, предъявляемых к системам менеджмента качества организаций. Независимое подтверждение соответствия системы менеджмента качества организации требованиям ISO 9001 может быть получено от третьей стороны - органа по сертификации.

Стандарты ISO серии 9000 состоят из:

ISO 9000 – Основные положения и словарь. Стандарт предназначен для установления единого понимания терминов и понятий, используемых в стандартах ISO серии 9000.

ISO 9001 – Требования. Стандарт определяет требования, которым должны соответствовать организации, желающие соответствовать стандарту.

Сторонние органы по сертификации предоставляют независимое подтверждение того, что организации соответствуют требованиям ISO 9001. Более миллиона организаций по всему миру прошли независимую сертификацию, что делает ISO 9001 одним из наиболее широко используемых инструментов управления в современном мире.

ISO 9004 – Рекомендации по улучшению деятельности. Стандарт содержит базирующиеся на восьми принципах менеджмента качества рекомендации, которые были разработаны для менеджеров для улучшения деятельности организации и удовлетворения требований всех заинтересованных сторон.

TickIT

TickIT – это схема сертификации систем качества для программного обеспечения, предложенная группой ведущих фирм и некоммерческих организаций Великобритании, работающих в области информатики. Контроль и спонсирование схемы осуществляются DTI. TickIT базируется на стандарте ISO 9001:1994. Таким образом, предметом TickIT является менеджмент предприятий, разрабатывающих программное обеспечение.

TickIT также включает в себя руководство. Кроме того, он предоставляет подробную информацию о том, как внедрить систему качества, а также об ожидаемой структуре и содержании, относящихся к действиям, связанным с программным обеспечением.

Часть А: Введение в TickIT и процесс сертификации.

Часть В: Руководство для заказчиков.

Часть С: Руководство для поставщиков.

Часть D: Руководство для аудиторов.

Часть E: Требования к системе управления качеством программного обеспечения.

Часть F: Требования к системе управления качеством программного обеспечения – перспектива процесса.

Стандарты SEI SW-CMM

Модель улучшения внутренних процессов SEI SW-CMM, разработанная Институтом программной инженерии университета Карнеги-Меллона, предназначена

исключительно для организаций, занимающихся разработкой ПО. Основная цель модели – постоянное самосовершенствование процессов разработки.

К преимуществам модели SEI SW-CMM относится ориентированность на организации, занимающиеся разработкой программного обеспечения. В модели более детально проработаны требования, специфичные для процессов разработки ПО. Вследствие этого в SEI SW-CMM приведены не только требования к процессам организации, но и примеры реализации этих требований.

Стандарт UML

Унифицированный язык моделирования (UML) – это язык моделирования общего назначения для разработки в области разработки программного обеспечения, который предназначен для предоставления стандартного способа визуализации дизайна системы. UML был создан ведущими экспертами в области объектно-ориентированного анализа и проектирования программных систем из корпорации Rational Software.

Первоначально создание UML было мотивировано желанием стандартизировать разрозненные системы обозначений и подходы к проектированию программного обеспечения.

В 1997 году UML был принят в качестве стандарта консорциумом Object Management Group(OMG), который занимается стандартизацией объектных технологий и с тех пор находится под управлением этой организации. На данный момент специалисты OMG занимаются всеми вопросами разработки языка UML.

Язык UML предназначен для визуального построения моделей программных систем. Графические модели UML с помощью соответствующих программных средств переводятся в программный код конкретной среды разработки.

Единая система программной документации (ГОСТ 19)

ЕСПД (единая система программной документации) – комплекс государственных стандартов (ГОСТ), которые устанавливают взаимосвязь между правилами разработки, оформления и обращения программ и программной документации. Стандарты ЕСПД определяют требования, регламентирующие разработку, сопровождение, изготовление и эксплуатацию программ.

Сейчас этот комплекс представляет собой систему межгосударственных стандартов стран СНГ (ГОСТ), действующих на территории Российской Федерации на основе межгосударственного соглашения по стандартизации.

Стандарты ЕСПД (как и другие ГОСТы) подразделяют на группы:

- общие положения;
- основополагающие стандарты;
- правила выполнения документации;
- правила выполнения эксплуатационной документации;
- правила обращения программной документации;
- резервные группы;
- прочие стандарты.

Из всех стандартов ЕСПД остановимся только на тех, которые могут чаще использоваться на практике:

ГОСТ (СТ СЭВ) 19.201-78 (1626-79). ЕСПД. Техническое задание. Требование к содержанию и оформлению.

ГОСТ (СТ СЭВ) 19.101-77 (1626-79). ЕСПД. Виды программ и программных документов (Переиздан в ноябре 1987г с изм.1). устанавливает виды программ и

программных документов для вычислительных машин, комплексов и систем независимо от их назначения и области применения.

ГОСТ 19.102-77. ЕСПД. Стадии разработки устанавливает стадии разработки программ и программной документации для вычислительных машин, комплексов и систем независимо от их назначения и области применения

ГОСТ 19.103-77 ЕСПД. Обозначение программ и программных документов

ГОСТ 19.106-78 ЕСПД. Требования к программным документам, выполненным печатным способом

ГОСТ 19.402-78 ЕСПД. Описание программы

ГОСТ 19781-90 Обеспечение систем обработки информации программное. Термины и определения.

Государственные стандарты РФ (ГОСТ Р)

В РФ действует ряд стандартов в части документирования ПС, разработанных на основе прямого применения международных стандартов ИСО.

ГОСТ Р ИСО/МЭК 9294-93 Информационная технология. Руководство по управлению документированием программного обеспечения. Целью стандарта является оказание помощи в определении стратегии документирования ПС; выборе стандартов по документированию; выборе процедур документирования; определении необходимых ресурсов; составлении планов документирования.

ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению. Стандарт определяет шесть комплексных характеристик, которые с минимальным дублированием описывают качество ПС (ПО, программной продукции): функциональные возможности; надежность; практичность; эффективность; сопровождаемость; мобильность.

ГОСТ Р ИСО 9127-94 Системы обработки информации. Документация пользователя и информация на упаковке для потребительских программных пакетов. Под документацией пользователя понимается документация, которая обеспечивает конечного пользователя информацией по установке и эксплуатации ПП. Под информацией на упаковке понимают информацию, воспроизводимую на внешней упаковке ПП.

ГОСТ Р ИСО/МЭК 8631-94 Информационная технология. Программные конструктивы и условные обозначения для их представления. Описывает представление процедурных алгоритмов.

Выводы

Подводя итоги хочется еще раз отметить, что использование стандартов при разработке программного обеспечения является очень важным пунктом, поскольку помогает лучше понять, как правильно выстроить и связать между собой функции, выполнение которых должно будет осуществляться ПО, как обеспечить соответствие требованиям, заявленному уровню качества, а также какая документация будет необходима для успешного создания и внедрения продукта.

Литература

1. Евгений Чернов [Электронный ресурс] – TickIT – схема сертификации систем качества для ПО – Режим доступа: <https://www.itweek.ru/themes/detail.php?ID=48313>

2. КРМС Менеджмент качества [Электронный ресурс] – Стандарты ИСО серии 9000 – Режим доступа: <https://www.kpms.ru/standart.htm>
3. Консалтинговая компания iTeam [Электронный ресурс] – Стандарты в области разработки программного обеспечения – Режим доступа: <https://blog.iteam.ru/standarty-v-oblasti-razrabotki-programmnogo-obespecheniya/amp/>
4. Анатолий Мурашов [Электронный ресурс] – Стандарты управления проектами – Режим доступа: <https://zaochnik.com/spravochnik/menedzhment/standarty-upravlenija-proektami/>
5. Бедрина С.Л. Разработка и стандартизация программного обеспечения: Учебное пособие. – Владивосток: Изд-во ВГУЭС, 2007. – 180 с.
6. Благодатских В.А. Волнин В.А., Посакалов К.Ф. Стандартизация разработки программных средств. – М.: Финансы и статистика, 2003.
7. Благодатских В.А. Экономика, разработка и использование программного обеспечения ЭВМ. – М.: Финансы и статистика, 1995.
8. Брауде Э.Дж. Технология разработки программного обеспечения: Пер. с англ. – СПб.: Питер, 2004.
9. Вигерс К. Разработка требований к программному обеспечению: Пер. с англ. М.: Русская редакция, 2004.
10. Вендеров А.М. Проектирование программного обеспечения экономических информационных систем – М.: Финансы и статистика, 2002.

УДК 004.9

ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ОБНАРУЖЕНИЯ ОБЛАСТИ ГЛАЗ НА ФОТОГРАФИИ ЛИЦА ЧЕЛОВЕКА

DESIGN AND DEVELOPMENT OF INFORMATION SYSTEM FOR DETECTING EYES AREAS AT HUMAN FACE PHOTO

Бурдуковский С.О.,
ФГБОУ ВО «Новосибирский государственный университет экономики и управления
«НИНХ»,
г. Новосибирск, Российская Федерация

S.O. Burdukowsky,
FSBEI HE “Novosibirsk State University of Economics and Management
“NINE”,
Novosibirsk, Russian Federation

e-mail: burdukowsky_stas@mail.ru

Аннотация. В данной статье автор описывает проектирование и разработку информационной системы обнаружения области глаз на фотографии лица человека. Описанная в статье работа является частью научного исследования, посвященного разработке информационной системы автоматического диагностирования косоглазия. Для разработанной информационной системы была выбрана клиент-серверная архитектура. Разработано три модуля информационной системы: серверная часть, клиентская часть для администратора и клиентская часть для пациента. Серверная часть является RESTful веб-API сервисом, а клиентские части представляют собой веб-приложения типа single page application (SPA). Для разработки использовались

современные, открытые, активно поддерживаемые технологии и подходы, что позволяет использовать в приложениях все актуальные возможности платформ и облегчает поддержку и дальнейшую разработку системы. Для веб-API был использован микро-фреймворк Flask, а для клиентских частей – фреймворк Angular версии 13. Использование клиент-серверной архитектуры позволяет в перспективе осуществить разработку дополнительных клиентов, например, для мобильных операционных систем iOS и Android. Данная статья будет полезна для исследователей и разработчиков, которые занимаются или планируют заниматься разработкой сложных систем, включающих пользовательские интерфейсы и вычисления на серверной стороне.

Abstract. In this article, the author describes the design and development of an information system for detecting for detecting eyes areas at human face photo. The work described in the article is part of a scientific study on the development of an information system for automatic diagnosis of strabismus. Developed information system has a client-server architecture. Three modules of the information system have been developed: the server part, the client part for the administrator and the client part for the patient. The backend is a RESTful web API service, and the frontend parts are a single page application (SPA). Modern, open, actively supported technologies and approaches were used in applications for development, which makes it possible to use all the relevant features of the goal platforms and facilitates the support and further development of the system. The Flask micro-framework was used for the web API, and the Angular framework (version 13) was used for the client parts. The using of client-server architecture makes it possible to develop additional clients in the future, for example, for iOS and Android mobile operating systems. This article will be useful for researchers and developers who are engaged or plan to develop complex systems that include user interfaces and server-side computing.

Ключевые слова: клиент-серверная архитектура, обнаружение объектов, API, фронтенд, бэкенд.

Keywords: client-server architecture, object detection, API, frontend, backend.

Описанная в статье работа является частью научного исследования, посвященного разработке информационной системы автоматического диагностирования косоглазия. Первый этап при обработке входного изображения – выделение областей с глазами. Эта статья посвящена разработке информационной системы обнаружения областей с глазами на фотографии лица человека.

Для разработанной информационной системы была выбрана клиент-серверная архитектура, так как она предполагает отдельную разработку серверной и клиентской частей, что позволяет в перспективе разработку дополнительных клиентов. Например, в данной статье описывается проектирование и разработка веб-клиентов информационной системы, а в будущем возможно появление мобильных клиентов для операционных систем iOS и Android.

Всего на данный момент разработано три модуля информационной системы:

- серверная часть,
- клиентская часть (административная панель),
- клиентская часть (интерфейс пациента).

Серверная часть является RESTful веб-сервисом. Она принимает HTTP-запросы от клиентов, обрабатывает их, работает с базой данных и формирует ответ в формате JSON.

Клиентские части являются одностраничными веб-приложениями (single page application или SPA), располагаемыми на отдельных HTTP веб-серверах и состоящими только из статичных файлов (HTML, JS, CSS, JSON, JPEG, PNG, GIF, SVG, файлы шрифтов).

REST (сокр. от англ. Representational State Transfer – «передача состояния представления») – архитектурный стиль взаимодействия компонентов распределённого приложения в сети. REST представляет собой согласованный набор ограничений, учитываемых при проектировании распределённой гипермедиа-системы. В определённых случаях (интернет-магазины, поисковые системы, прочие системы, основанные на данных) это приводит к повышению производительности и упрощению архитектуры. В широком смысле компоненты в REST взаимодействуют наподобие взаимодействия клиентов и серверов во Всемирной паутине. В сети Интернет вызов удалённой процедуры может представлять собой обычный HTTP-запрос (обычно «GET» или «POST»); такой запрос называют «REST-запрос»), а необходимые данные передаются в качестве параметров запроса [1].

Такое разделение позволяет распределить нагрузку на веб-приложение в целом. Также такой подход позволит значительно уменьшить сетевой трафик, так как все файлы кроме одной страницы HTML будут кэшироваться клиентским браузером, а взаимодействие с сервером будет осуществляться путем передачи легковесного JSON. Этот подход позволяет значительно снизить нагрузку с серверной части, так как ей уже не нужно отдавать статику и формировать HTML страницы для пользователя. Серверная часть будет только работать с базой данных, осуществлять необходимые вычисления и формировать ответ.

Для разработки серверной части был выбран веб-фреймворк Flask. Выбор был сделан в пользу этого веб-фреймворка, потому что он имеет открытый исходный код, активно поддерживается и имеет большое сообщество разработчиков. Также этот микрофреймворк имеет очень маленький размер и сознательно предоставляет лишь самые базовые возможности, что обеспечивает высокую производительность работы [2].

В качестве языка программирования выбран Python, так как модуль автоматического детектирования области глаз также разрабатывается на этом языке программирования с использованием платформы TensorFlow, и выбор Python позволит объединить эти модули в одну информационную систему.

API помимо того, что служит для обработки HTTP-запросов от клиентских приложений с одной стороны и для вызова модуля детектирования области глаз с другой, еще производит журналирование HTTP-запросов и сохранение результатов работы модуля обнаружения в базе данных PostgreSQL.

Разработанный API содержит два контроллера обработки HTTP-запросов. Первый – MainController – обрабатывает запросы от клиентского приложения пациента, в котором пользователь загружает фотографию глаз для дальнейшего анализа. Второй – AdminController – обрабатывает запросы, которые осуществляет клиентское приложение администратора информационной системы. Эти запросы служат для предоставления информации о версии API, версии модуля обнаружения, просмотра журналов использования API и просмотра информации о результатах обнаружения, сохраненных в базе данных PostgreSQL.

MainController предоставляет POST-запрос «recognize», который принимает фотографию глаз человека. При обработке запроса принятая фотография отправляется модулю обнаружения. В зависимости от статуса работы модуля обнаружения формируется ответ в формате JSON, который содержит статус обработки запроса («успешно» или «не успешно»), информацию о координатах областей глаз. Если сервис распознавания не смог обнаружить глаза на фотографии, то возвращается

соответствующее сообщение и пользователю нужно будет повторить отправку. После обработки запроса выполняется сохранение результата работы модуля обнаружения и текущее время в базу данных.

AdminController предоставляет три GET-запроса: «version», «log» и «results». Запрос «version» возвращает клиенту ответ в формате JSON, который содержит объект с двумя полями: «api» и «recognitionSystem». В полях указываются версии API и модуля обнаружения соответственно. Запрос «log» возвращает клиенту журнал HTTP-запросов к API. Запрос имеет два параметра – «from» и «limit», которые позволяют запросить журнал от любой даты и ограничить количество записей. Ответ на запрос «log» представляет собой JSON-массив с сущностями, описывающими запросы (поля «address», «timestamp» и «ip», содержащие адрес, время запроса и IP-адрес клиента соответственно). Запрос «results» возвращает клиенту результаты работы модуля обнаружения. Этот запрос, подобно запросу «log», имеет два параметра – «from» и «limit», которые позволяют запросить результаты от любой даты и ограничить количество записей. Ответ на запрос «results» представляет собой JSON-массив с сущностями, описывающими результат работы модуля распознавания (поля «response» и «timestamp» и «status», содержащие тело ответа модуля обнаружения, время проведения обнаружения и информацию об успехе обработки фотографии соответственно).

Для разработки клиентских частей был выбран фреймворк Angular версии 13, как открытый, хорошо поддерживаемый, имеющий большое сообщество разработчиков инструмент разработки пользовательских интерфейсов. Angular позволяет разрабатывать сложные информационные системы корпоративного уровня, предоставляет модульную архитектуру, поддерживает все современные браузеры. Фреймворк разработан компанией Google [3], постоянно обновляется и поддерживается, что подчеркивает его актуальность. В качестве языка программирования используется TypeScript, разработанный компанией Microsoft. TypeScript является обратно совместимым с JavaScript и компилируется в последний. TypeScript имеет строгую типизацию, что является решающим фактором при выборе этого инструмента.

Архитектура клиентских частей имеет общие черты, а именно представляет собой три модуля: AppModule, CoreModule и SharedModule. AppModule включает в себя компоненты, отвечающие за работу страниц веб-интерфейса. CoreModule включает в себя сервисы для работы с серверной частью и хранилищем данных. А SharedModule включает в себя компоненты и директивы общего назначения, которые могут использоваться в различных submodule AppModule.

Веб-интерфейс пациента представляет собой главный экран, экран диагностики и экран результатов диагностирования. Каждый экран со стороны архитектуры приложения представляет собой отдельный ViewModule, включенный в AppModule приложения.

На главном экране отображается приветственное сообщение и описание информационной системы. Также на этом экране имеется переход к странице диагностики.

На странице диагностики веб-приложение запрашивает браузерное разрешение на использование веб-камеры, в случае если пациент хочет сделать фотографию глаз прямо в веб-интерфейсе. Также пациент может загрузить с устройства уже готовую фотографию. На фотографии должны быть четко видны глаза. После выбора фотографии с устройства или создания снимка в браузере, картинка отправляется серверной части путем HTTP-запроса. Сервер обрабатывает изображение и в ответ на запрос выдает информацию в формате JSON, которая содержит статус обработки запроса («успешно» или «не успешно»), информацию о координатах обнаруженных объектов. Если сервис

распознавания не смог обнаружить глаза на фотографии, то возвращается соответствующее сообщение и пользователю нужно будет повторить отправку.

После успешной отработки веб-интерфейс переходит на экран результатов. На этом экране пользователь может увидеть выделенные области с глазами на изображении.

Веб-интерфейс администратора представляет собой экран входа, экран просмотра журнала работы API и экран просмотра результатов работы модуля обнаружения.

Экран входа представляет собой форму входа с полями «E-mail» и «Пароль». Если пользователь введет корректные данные входа и отправит форму, то будет произведен переход к основному экрану. В случае неверных данных входа, пользователь увидит соответствующее сообщение.

Основной экран представляет собой боковое меню с двумя элементами: «Журнал» и «Результаты». По умолчанию открывается страница «Журнал». На ней администратор может просматривать журнал работы API. Изначально запрашивается 100 записей от текущей даты. При прокрутке содержимого страницы до конца, автоматически происходит запрос следующих 100 записей от даты последней загруженной записи.

На странице «Результаты» администратор может просматривать результаты работы модуля обнаружения, которые содержат тело ответа модуля, время проведения обнаружения и информацию об успехе обработки фотографии соответственно. Изначально запрашивается 20 записей, а запрос последующих записей осуществляется таким же способом, как и на странице «Журнал».

Выводы

В ходе работы была разработана информационная система обнаружения области глаз на фотографии лица человека, имеющая клиент-серверную архитектуру и состоящая из трех модулей:

- серверная часть,
- клиентская часть (административная панель),
- клиентская часть (интерфейс пациента).

Серверная часть является RESTful веб-API сервисом, обслуживающим запросы клиентских частей, а клиентские части являются веб-приложениями для администраторов и пациентов. Благодаря клиент-серверной архитектуре разработанная система готова к расширению – разработке мобильных клиентов для различных операционных систем. Для разработки использовались современные, открытые, активно поддерживаемые технологии и подходы, что позволяет использовать в приложениях все актуальные возможности платформ и облегчает поддержку и дальнейшую разработку системы.

Литература

1. REST – Википедия [Электронный ресурс]. – URL:<https://ru.wikipedia.org/wiki/REST> (дата обращения 17.04.22).
2. Flask (веб-фреймворк) – Википедия [Электронный ресурс]. – URL:[https://ru.wikipedia.org/wiki/Flask_\(веб-фреймворк\)](https://ru.wikipedia.org/wiki/Flask_(веб-фреймворк)) (дата обращения 17.04.22).
3. Angular (фреймворк) – Википедия [Электронный ресурс]. – URL:[https://ru.wikipedia.org/wiki/Angular_\(фреймворк\)](https://ru.wikipedia.org/wiki/Angular_(фреймворк)) (дата обращения 17.04.22).

УДК 004.031.4

**ИССЛЕДОВАНИЕ КАЧЕСТВА ДАННЫХ В ГРАЖДАНСКОЙ НАУКЕ
С ПОМОЩЬЮ РЕШЕНИЙ INDUSTRY 4.0****DATA QUALITY INVESTIGATION IN CITIZEN SCIENCE
WITH THE HELP OF INDUSTRY 4.0 SOLUTIONS**

Вердиева Н.Н.,
Институт Информационных Технологий НАНА,
г. Баку, Азербайджан

N.N. Verdiyeva,
Institute of Information Technology of ANAS,
Baku, Azerbaijan

e-mail: nergiz_verdieva@mail.ru

Аннотация. В наши дни применение современных информационных технологий распространилось на все области жизни граждан. Наука, являющаяся одной из этих областей, стала развиваться с большей скоростью, постепенно превращаясь из традиционной науки в э-науку. Это, в свою очередь, послужило толчком для развития различных направлений э-науки, в том числе гражданской науки. В современном мире информационные технологии широко применяются во всех сферах жизни, включая научные исследования. Гражданская наука, которая характеризует участие волонтеров в процессе научных исследований, также стала развиваться намного быстрее за счет применения современных технологий и последних достижений в области ИКТ. В период четвертой промышленной революции, называемой также Industry 4.0, решение различных интеллектуальных задач значительно облегчилось, в том числе сократилось время их выполнения за счет использования технологий Industry 4.0. Эти технологии оптимизируют информационные процессы, тем самым улучшая качество научных исследований. Статья посвящена исследованию качества данных в гражданской науке с помощью решений Industry 4.0. Рассмотрены работы разных авторов на тему качества данных в проектах гражданской науки. Обсуждены вопросы верификации и валидации данных, сгенерированных гражданами-учеными в проектах гражданской науки. Показаны различные формы верификации данных, а также их применение на разных этапах проектов гражданской науки. Рассмотрены основные стратегии обеспечения и контроля качества данных в проектах гражданской науки.

Abstract. Nowadays, the use of modern information technologies has spread to all areas of life of citizens. Science, which is one of these areas, began to develop at a faster pace, gradually transforming from traditional science into e-science. This, in turn, served as an impetus for the development of various areas of e-science, including citizen science. In the modern world, information technologies are widely used in all spheres of life, including scientific research. Citizen science, which characterizes the participation of volunteers in the process of scientific research, has also developed much faster due to the use of modern technologies and the latest advances in ICT. During the period of the fourth industrial revolution, also called Industry 4.0, the solution of various intellectual tasks was greatly facilitated, including the reduction of their execution time through the use of Industry 4.0 technologies. These technologies optimize information processes, thereby improving the

quality of scientific research. The article is dedicated to data quality investigation in citizen science with the help of Industry 4.0 solutions. The works of different authors on the topic of data quality in citizen science projects are considered. Issues of verification and validation of data generated by citizen scientists in citizen science projects were discussed. Various forms of data verification are shown, as well as their application at different stages of citizen science projects. The main strategies for ensuring and controlling the quality of data in citizen science projects are considered.

Ключевые слова: э-наука, гражданская наука, качество данных, Industry 4.0, ИКТ.

Keywords: e-science, citizen science, data quality, Industry 4.0, ICT.

В современном мире информационные технологии широко применяются во всех сферах жизни, включая научные исследования. Гражданская наука (ГН), которая характеризует участие волонтеров в процессе научных исследований, также стала развиваться намного быстрее за счет применения современных технологий и последних достижений в области ИКТ.

В период четвертой промышленной революции, называемой также Industry 4.0, решение различных интеллектуальных задач значительно облегчилось, в том числе сократилось время их выполнения за счет использования технологий Industry 4.0. Эти технологии оптимизируют информационные процессы, тем самым улучшая качество научных исследований. Применение инструментов Industry 4.0 в ГН поможет гражданам-ученым в процессах сбора, обработки, хранения и передачи данных.

Сбор данных является основной целью многих проектов ГН. Успешные проекты должны обеспечивать качество, полезность и защиту данных.

Качество данных – это метрика, указывающая, в какой степени набор данных (dataset) соответствует конкретной цели.

Измерение качества данных основано на таких характеристиках, как точность, полнота, надежность, согласованность и актуальность.

Вопрос качества данных в проектах ГН остается важным моментом, который необходимо должным образом контролировать. Некоторые исследования показывают, что результаты, собранные гражданами, более изменчивы, чем результаты, собранные профессиональными учеными, в то время как другие исследования показывают, что качество данных в обеих группах одинаково. В любом случае мониторинг качества данных должен быть главной целью проекта.

Для повышения точности данных в проектах ГН необходимо принять следующие меры:

- сбор справочных данных в начале проекта;
- предоставление возможности гражданам участвовать в процессах мониторинга;
- обеспечение предварительного обучения гражданских ученых и проверки проводимых тренингов;
- увеличение количества участников и объема собираемых данных;
- получение дополнительной информации о профилях участников для адаптации протоколов к знаниям и навыкам группы граждан, привлекаемых к исследованию.

Актуальность проблемы качества данных в проектах ГН можно наблюдать по научным работам в этой сфере. Так, в литературе [1] авторы представили широкий спектр методов для оценки и улучшения качества данных. Из-за разнообразия и сложности этих методов, исследования в последнее время были сосредоточены на определении методологий, которые помогают выбирать, настраивать и применять

методы оценки и улучшения качества данных. В статье представлено систематическое и сравнительное описание таких методологий. Методологии сравниваются по нескольким параметрам, включая методологические фазы и этапы, стратегии и техники, параметры качества данных, типы данных и, наконец, типы информационных систем, для которых предназначена каждая методология.

Авторы статьи [2] обсуждают результаты опроса о механизмах, используемых проектами ГН для обеспечения качества научных исследований. В ходе обзора проектов ГН обнаружено, что в большинстве проектов используется несколько механизмов для обеспечения качества данных и соответствующих уровней проверки. Они создали структуру из 18 механизмов, обычно используемых в проектах ГН для обеспечения качества данных, на основе непосредственного опыта авторов и обзора данных опроса, отметив две категории источников ошибок (протоколы, участники) и три возможных точки вмешательства (до, во время и после участия), которые можно использовать для разработки проекта.

В работе [3] рассмотрена гибкая система оценки качества краудсорсинговых данных. Здесь описывается структура для квалификации геолокационных данных, собранных из неавторитетных источников, которая позволяет проводить оценку для конкретных тематических исследований путем создания рабочего процесса, поддерживаемого онтологическим описанием ряда вариантов:

- 1) позиционирование сервисов на основе местоположения перенаправляет пользователей в области, представляющие интерес для организаторов проекта;
- 2) очистка данных удаляет ошибочные записи;
- 3) автоматическая проверка выполняет предварительную проверку достоверности собранных данных;
- 4) сравнение с достоверными данными повышает достоверность данных;
- 5) проверка на основе моделей сравнивает данные с данными из моделей или ранее проверенными данными из краудсорсинга;
- 6) анализ связанных данных сочетает в себе множество свободно доступных данных (Big Data) и соответствующие методы интеллектуального анализа данных для обеспечения достоверности и качества данных;
- 7) семантическая гармонизация преобразует входные данные для обеспечения соответствия или обогащения онтологии.

В работе [4] собраны данные, предоставленные iNaturalist, платформой ГН, на которой размещаются проекты по сбору данных о биоразнообразии, где добровольцы участвуют в наблюдениях за образцами по всему миру и вносят предложения по идентификации таксонов (ID) в любой записи.

Оценка качества, предоставленная iNaturalist, использовалась для определения классов, которые указывают, было ли наблюдение качественным или нет, поскольку в настоящей работе применяется контролируемое обучение.

Метрика iNaturalist определяет три категории качества:

- случайная – это начальный уровень, но его также можно назначить для наблюдений, информация о которых была подвергнута сомнению;
- идентификатор потребности указывает, что наблюдение имеет дату, географическую привязку, фото или звук и не относится к человеку, но нет идентификация вида;
- исследовательская степень применяется, когда пользователи соглашаются в отношении идентификации. Этот уровень используется для указания на наличие качества, а два других – на его отсутствие.

В одной из глав книги «The Science of Citizen Science» [5] обсуждается обширная и сложная тема качества данных в ГН, которая считается спорной сферой, поскольку

разные проекты и заинтересованные стороны стремятся к разным уровням точности данных. Здесь рассмотрено, как обеспечить надежность данных, собранных гражданскими учеными.

Авторы представили общий обзор основных тем и проблем, связанных с качеством данных в ГН, механизмов обеспечения и повышения качества, а также некоторые выводы о передовом опыте и дальнейших действиях. Они призывают проекты ГН делиться информацией о своих неудачах в работе с данными. Показано, как обеспечение качества данных придает доверие и устойчивость проектам ГН.

Авторы работы [6], признавая необходимость повышения качества данных ГН, описывают вопросы оценки и контроля качества (ОК/КК) данных и предлагают разные взгляды на аспекты обеспечения качества данных ГН и проведения исследований по этим вопросам. Здесь показаны четыре параметра качества, которые возникают на протяжении жизненного цикла данных, проблемы, направления и подходы к повышению качества данных ГН. Представлен краткий обзор возможностей повышения качества данных ГН при наборе, отборе, самоотборе и обучении граждан-ученых. Описываются отдельные вопросы, относящиеся к прозрачности информации о методах ОК/КК при подготовке данных, а также важность документирования качества данных ГН. Показана необходимость создания рубрик для оценки уровней качества данных. Предложены рекомендации по улучшению качества данных в ГН.

Данные в ГН должны пройти верификацию и валидацию, прежде чем их можно будет использовать.

Для верификации данных используются четыре подхода:

- коллегальная верификация,
- экспертная верификация,
- автоматическая оценка качества
- оценка качества на основе моделей.

Мы рассмотрим последние два из них, чтобы показать применение решений Industry 4.0 в данном контексте.

Автоматическая оценка качества предполагает использование программных систем для автоматической оценки качества данных, сгенерированных в рамках проекта ГН. Существует широкий спектр подходов, таких как алгоритмы интеллектуального анализа данных, которые фильтруют и ищут проблемные данные, статистический анализ (правдоподобие данных) и системы квалификации. По мере того, как подходы к искусственному интеллекту (ИИ) становятся все более изощренными и все более доступными в программном обеспечении, их можно использовать для проведения более ресурсоемких автоматизированных оценок качества.

Оценка качества на основе моделей выходит за рамки автоматических методов фильтрации, которые могут учитывать случайные изменения, и устраняет остаточные ошибки, используя явную модель того, как ожидается, что интересующее явление будет изменяться в пространстве или во времени. Это требует конкретного понимания того, как ведут себя соответствующие явления, и нужны соответствующие специалисты.

Этот подход может быть более эффективным при установлении статистической релевантности результатов, а также экстремальных или неожиданных значений в данных. Примеры включают процедуры устранения предвзятости и составление рейтингов участников на основе выявленных источников систематических ошибок в архиве наблюдений [5].

Обеспечение и контроль качества данных в ГН могут осуществляться с помощью двух стратегий:

- 1) обеспечивающая стратегия, включающая набор действий, обеспечивающих качество данных ГН до определенного уровня,

или

2) контролирующая стратегия, включающая набор действий, которые контролируют качество ГН и извлекают уроки из предыдущих неудач.

Для обеспечения качества данных требуется набор критериев, которые заблаговременно ограничивают ввод данных, таких как:

– профилирование, которое оценивает сборщиков данных, чтобы понять проблемы качества, включая влияние неопределенности в данных и то, как ее можно зафиксировать или отследить.

– предварительное тестирование включает в себя сбор выборочных данных до начала проекта ГН с участием как экспертов, так и новичков. Это может помочь выявить непредвиденные источники ошибок или других проблем, которые можно исправить до начала проекта.

– стандартизация гарантирует, что ожидаемые данные соответствуют правилам качества и схемам, относящимся к предметной области.

– инструменты оперативного исправления или очистки данных позволяют автоматически исправлять некоторые ошибки перед отчетом, например, автокоррекция геокодирования адресных данных, проверки топологии и принудительный выбор значения атрибута из списка словаря.

– сопоставление или связывание облегчает выравнивание или объединение похожих записей данных, что помогает избежать избыточности данных.

Контроль качества данных включает в себя набор действий, позволяющих контролировать качество данных после запуска проекта, например:

– триангуляция, представляющая собой комбинацию нескольких критериев и методов для обеспечения качества данных.

– рекурсивный мониторинг отслеживает качество данных и генерирует отчеты о неопределенностях и отклонениях. Эти отчеты можно использовать для поддержания или улучшения качества данных, а также для предоставления отзывов о ходе разработки проекта.

– в результате обучения участники лучше понимают качество данных и способны оценить минимальные требования к качеству данных для каждого проекта ГН.

– соблюдаются протоколы и стандарты соответствия для обеспечения согласованности и однородности собранных данных. Использование протоколов и стандартов не должно отрицательно сказываться на уровне занятости гражданских ученых.

– взаимно совместимые информационные системы обеспечивают долгосрочное хранение, курирование и архивирование данных проектов ГН.

– использование международных стандартов, таких как ISO19115, ISO19157 и ISO8000, рекомендуется в качестве ориентира для контроля качества проектов ГН.

– сбор и публикация данных в соответствии с принципами открытой науки и лицензиями открытого доступа, которые следуют принципам FAIR (находимость, доступность, совместимость, повторное использование). Это обеспечивает неограниченный доступ к данным и позволяет повторно использовать данные. Использование принципов FAIR максимизирует ценность данных.

– регистрация и отчетность по методам обеспечения качества, так как описания методов обеспечения качества ГН часто отсутствуют. Эта информация должна быть предоставлена в описании или метаданных проекта, или набора данных, чтобы можно было избежать подобных сбоях в будущем.

Выводы

Исследование показало, что в дополнение к установлению достоверности и доверия, обмен информацией о методах обеспечения качества данных в проектах ГН может помочь сотрудничеству путем выявления общих проблем.

Обеспечение качества данных в проектах ГН с помощью технологий Industry 4.0 может улучшить репутацию проекта, обеспечить широкое использование результатов и способствовать долгосрочной устойчивости проектов, а также вовлечению граждан в проекты ГН.

Литература

1. Batini C., Cappiello C., Francalanci C., Maurino A. Methodologies for Data Quality Assessment and Improvement//ACM Computing Surveys. 2009. DOI: 10.1145/1541880.1541883/
2. Wiggins A., Newman G., Stevenson R.D., Crowston K. Mechanisms for Data Quality and Validation in Citizen Science//IEEE 7th International Conference on E-Science: Workshop Proceedings. December 5-8, 2011. Stockholm, Sweden. DOI:10.1109/eScienceW.2011.27
3. Meek S., Jackson M., Leibovici D. A flexible framework for assessing the quality of crowdsourced data//AGILE'2014 International Conference on Geographic Information Science. Castellon, Spain. 2014.
4. Leocadio J.N., Saraiva A.M. A data-driven study of citizen science data quality assessment profile // 20th International Conferences on WWW/Internet 2021 and Applied Computing 2021. P. 93-100.
5. Balázs B., Mooney P., Nováková E., Bastin L., Jokar Arsanjani J. Data Quality in Citizen Science. In: Vohland K. et al. (eds) The Science of Citizen Science // Springer, Cham. 2021. https://doi.org/10.1007/978-3-030-58278-4_8.
6. Downs R.R., Ramapriyan H.K., Peng G., Wei Y. Perspectives on Citizen Science Data Quality // Frontiers in Climate. April 2021. <https://doi.org/10.3389/fclim.2021.615032>

УДК 004.93

АЛГОРИТМЫ ПРАКТИЧЕСКОЙ ПРИВЯЗКИ БПЛА К МЕСТНОСТИ

ALGORITHMS FOR THE PRACTICAL BINDING OF UAVS TO THE TERRITORY

Хайдошкина Ю.В., Наташкина Е.А.,
 ГАУ ТО «Центр информационных технологий»,
 г. Тула, Российская Федерация

Khaydoshkina Yu.V., Natashkina E.A.,
 SAI TR «Center of Information Technologies»,
 Tula, Russian Federation

e-mail: 5el@ro.ru

Аннотация. В данной статье рассматривается вопрос, связанный с практической стороной привязки полученных снимков с беспилотных летательных аппаратов (БПЛА)

к местности. Приводятся основные этапы технологического цикла технологии ГИС. Подробно разбирается этап интерпретации и формализации данных, полученных после проведения аэрофотосъемки местности с БПЛА. В статье отмечается существование различных алгоритмов, которые используются на практике для привязки БПЛА к местности, а также для фильтрации изображения, полученного в результате съемки, с целью создания образа, пригодного для совмещения с виртуальной моделью местности. Так, были приведены: фильтр Гаусса, который обеспечивает фильтрацию низких частот в спектре изображения, при этом оставляя доминирующее положение центрального пикселя, но принимая в учет значения соседних пикселей; и противоположная процедура обработки изображения, которая получила название деблюринг, направленная на повышение резкости и устранения искажений, вызываемых при движении и вибрации беспилотных летательных аппаратов. Рассматриваются алгоритмы определения контурных корреляторов, а именно алгоритм трассировки квадратов, алгоритм трассировки окрестностей Мура и алгоритм Тео Павлидиса. Они необходимы для выделения на первоначальном изображении контуров интересующих объектов.

Abstract. This article discusses the issue related to the practical side of linking the images obtained from unmanned aerial vehicles (UAVs) to the terrain. The main stages of the technological cycle of GIS technology are given. The stage of interpretation and formalization of the data obtained after aerial photography of the area from the UAV is analyzed in detail. The article notes the existence of various algorithms that are used in practice to bind the UAV to the terrain, as well as to filter the image obtained as a result of shooting, in order to create an image suitable for combining with a virtual terrain model. So, the following were given: a Gaussian filter, which provides low-pass filtering in the image spectrum, while leaving the dominant position of the central pixel, but taking into account the values of neighboring pixels; and the opposite image processing procedure, which was called deblurring, aimed at increasing sharpness and eliminating distortions caused by the movement and vibration of unmanned aerial vehicles. Algorithms for determining contour correlators are considered, namely the square tracing algorithm, the Moore neighborhood tracing algorithm and Theo Pavlidis algorithm. They are necessary to highlight the contours of the objects of interest in the initial image.

Ключевые слова: беспилотные летательные аппараты (БПЛА), информационные технологии, аэрофотосъемка, геоинформационная система, координаты местности, информатизация.

Keywords: unmanned aerial vehicles (UAVs), information technology, aerial photography, geoinformation system, terrain coordinates, informatization.

Основной задачей использования данных аэрофотосъемки, полученных с помощью БПЛА является их совмещение и сравнение с существующими в различных базах топографическими данными (в том числе с базой Единого государственного реестра недвижимости [5]). Современным методом решения этой задачи является разработка и внедрение так называемых геоинформационных систем (ГИС). ГИС – это современная компьютерная технология для картирования и анализа объектов реального мира [3].

Технология ГИС позволяет организовать полностью замкнутый технологический цикл, включающий следующие этапы:

- аэрофотосъемка местности с БПЛА;
- интерпретация и формализация полученных данных;

- сопоставление формализованных данных с имеющимися базами;
- проведение анализа, формирование отчета.

В данной статье рассмотрим второй этап технологического цикла.

Интерпретация и формализация данных, полученных в результате аэрофотосъемки с БПЛА, заключается в решении следующих задач:

- привязка полученных данных к координатам местности;
- фильтрация полученных изображений с целью минимизации шумов и артефактов съемки, а также четкого выделения контуров интересующих объектов;
- перевод данных в удобный цифровой формат.

Привязка данных к координатам на местности в настоящее время сложностей не представляет ввиду широкой доступности навигационных систем позиционирования (ГЛОНАСС, GPS). Для повышения точности на местности, где проводится аэрофотосъемка могут быть дополнительно размещены специальные маяки, координаты которых определены с высокой точностью геодезическими методами, и которые могут быть использованы для высокоточной привязки данных съемки к конкретной местности. В качестве таких маяков могут быть использованы как специально размещенные объекты, так и хорошо заметные объекты на фотографируемой местности.

Существуют различные алгоритмы для практической привязки БПЛА к местности, одним из популярных и широко используемых, является алгоритм ASIFT, позволяющий получить координаты центра снимка местности по сопоставлению с эталонными координатами и изображениями местности из имеющихся баз данных [4]. В источнике [4] также представлено подробное описание работы алгоритма. Следует учитывать, что существуют требования, которые применяются к определению координат и точек границ различных участков, а также зданий, сооружений и прочих объектов. Все они определены в приказе Федеральной службы государственной регистрации, кадастра и картографии.

После проведения привязки данных аэрофотосъемки к эталонной базе данных необходимо провести обработку полученного изображения для точного определения границ интересующих объектов. Для этого используются различные алгоритмы фильтрации полученного изображения и его совмещения с виртуальной моделью местности (ВММ). Целью данного этапа обработки является:

- генерация и совмещение ракурса ВММ по координатам бортовых систем БПЛА и формирование растрового изображения;
- устранение артефактов растрового изображения;
- фильтрация полученных контуров для повышения точности;
- выделение на изображении контуров интересующих объектов.

Применяются различные алгоритмы фильтрации изображения, полученного с БПЛА, для создания образа, пригодного для совмещения с ВММ [1], например:

1. Фильтр Гаусса. По сути представляет собой фильтр низких частот в спектре изображения, сохраняет доминирующее положение центрального пикселя, но учитывает значения соседних пикселей. Приводит к некоторому размытию изображения. Ниже приведена функция фильтра Гаусса для двух пространственных измерений:

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/(2\sigma^2)}$$

где u, v – координаты точки;

σ – среднеквадратическое отклонение нормального распределения.

$$r^2 = u^2 + v^2,$$

r^2 – двумерный радиус размытия.

В случае проведения двух измерений представленная формула позволяет задать поверхность, которая имеет вид концентрических окружностей с распределением Гаусса от точки в центре. К исходному изображению применимы матрицы свертки, которые строятся из пикселей с распределением отличных от нуля [6].

2. Деблюринг. Это процедура обработки изображения противоположная фильтру Гаусса, позволяет повысить резкость, устранить искажения, вызванные движением и вибрацией БПЛА. Деблюринг основан на расчете средневзвешенной функции соседних пикселей (окрестностей центрального пикселя). Простейший алгоритм приведен в формуле:

$$x=y + \alpha (y-B(y)) = (1+\alpha)y-\alpha B(y),$$

где B – оператор размытия;

y – размытое изображение;

x – обработанное резкое изображение.

Для четкого выделения на полученном изображении контуров интересующих объектов используются так называемые контурные корреляторы. Существует ряд алгоритмов, позволяющих их реализовать, например:

1. алгоритм трассировки квадратов. Основная идея этого алгоритма заключается в последовательном переходе от начального пикселя полученного изображения (выделяется достаточно произвольно) к последующим пикселям. При этом, если анализируемый пиксель «темный», то производится поворот налево, а если «светлый» – направо. Алгоритм продолжает действие, пока не происходит возврат на начальный пиксель. Все найденные в процессе работы алгоритма «темные» пиксели и составят контур анализируемого объекта. Очевидным недостатком данного метода является игнорирование алгоритмом пикселей, которые расположены по диагонали, что может привести к существенным искажениям при восстановлении контура. Также не всегда возможно возвращение в начальный пиксель, что приводит к «зависанию» алгоритма.

2. алгоритм трассировки окрестностей Мура. Окрестность Мура – это множество из 8 смежных пикселей, которые имеют с анализируемым пикселем одно общее ребро или вершину.

Графически это можно представить следующим образом (рисунок 1).

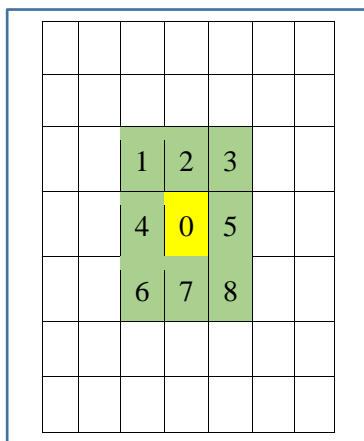


Рисунок 1. Окрестность Мура для пикселя 0

Цифрами 1-8 обозначена окрестность Мура пикселя, обозначенного цифрой 0.

Идея алгоритма заключается в том, что, начав движение от выбранного начального пикселя, двигаться по часовой стрелке. При этом при попадании на «темный» пиксель необходимо вернуться назад на предыдущий «светлый» пиксель и начинаем его обходить по окрестности Мура, пока не попадем на «темный пиксель». Завершение работы алгоритма происходит при возврате на начальный пиксель. Данный алгоритм лучше находит контуры объектов, но также имеет проблемы с точкой останова.

3. Алгоритм Тео Павлидиса.

Он лишен вышеупомянутых недостатков. Аналогично предыдущим алгоритмам достаточно произвольно определяется начальный пиксель, но с одним важным ограничением: левый соседний пиксель не «темный». Иначе говоря, вход в алгоритм начинается с «темного» пикселя, у которого левый пиксель – «светлый». После этого делается смещение на один пиксель вперед и на один пиксель влево. Если пиксель, на котором находится алгоритм окажется «темным» (рисунок 2), то текущим объявляется пиксель 1 и действие повторяется.

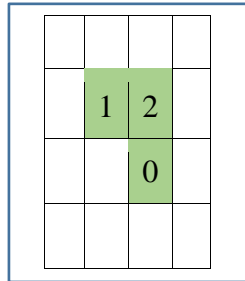


Рисунок 2. Первый шаг алгоритма Тео Павлидиса с попаданием на «темный» пиксель

Если же пиксель 1 окажется «светлым», то проверяется пиксель 2 (рисунок 3).

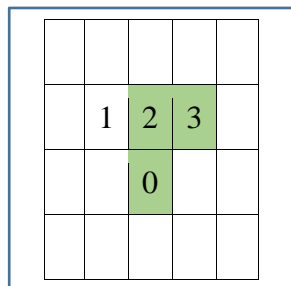


Рисунок 3. Первый шаг алгоритма Тео Павлидиса с попаданием на «светлый» пиксель

В этом случае текущим объявляется пиксель 2 и алгоритм движется на шаг вперед и попадает на пиксель 2, который объявляется текущим.

Третья возможная ситуация – когда пиксели 1 и 2 белые (рисунок 4).

В этом случае изменяется вектор движения алгоритма поворотом на 90 градусов вправо и действие повторяется.

Завершение работы алгоритма происходит, когда пиксели 1, 2 и 3 окажутся «светлыми».

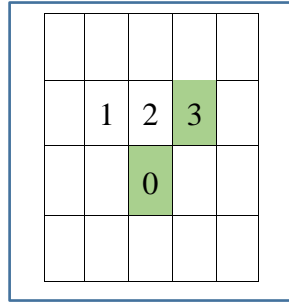


Рисунок 4. Первый шаг алгоритма Тео Павлидиса когда пиксели 1 и 2 «светлые»

Выводы

Рассмотренные алгоритмы позволяют выделить на первоначальном изображении контуры интересующих объектов. Поскольку координатная привязка к ВММ на данном этапе уже произведена, то можно переходить к сопоставлению найденных объектов к объектам, размещенным на ВММ. Наиболее эффективным способом такого сопоставления является использование нейросетей.

После выполнения всех предыдущих этапов, возможно проведение сопоставления полученных в результате съемки данных с эталонной базой данных с целью выявления расхождений. В настоящее время популярным технологическим решением для подобных задач являются сверточные нейронные сети [2]. Особенностью нейросети для обработки данных аэрофотосъемки является наличие функционала преобразования изображений (обычно обозначаются image-to-image или pix2pix), таким функционалом обладает программа PhotoScan, которая является частью комплекса для аэрофотосъемки с БПЛА GeoScan. Пример реализации такого функционала приведен на рисунке 5.



Рисунок 5. Пример преобразования фото со спутника в карту (изображение из открытых источников)

Таким образом, можно сделать вывод, что применение БПЛА открывает широкие возможности для специалистов в области землеустройства. Прежде всего использование

метода значительно увеличивает оперативность получения информации, снижая при этом затраты. Полученная в результате аэрофотосъемок информация обладает существенными особенностями. Достоверность и точность являются явными преимуществами использования данного метода на практике.

Литература

1. Акинин М.В., Соколова А.В., Никифоров М.Б. Нейросетевой метод оперативного совмещения данных аэрофотосъемки и виртуальной модели местности//Известия Тульского государственного университета. Технические науки. 2014. № 9-1. С. 106-114.
2. Багаев И.И. Анализ понятий нейронная сеть и сверточная нейронная сеть, обучение сверточной нейросети при помощи модуля TENSORFLOW//Математическое и программное обеспечение систем в промышленной и социальной сферах. 2020. Т. 8. № 1. С. 15-22.
3. Гусева А.В. Геоинформационные системы//Горный информационно-аналитический бюллетень (научно-технический журнал). 2013. № S5. С. 50-55.
4. Корнеев М.А., Максимов А.Н., Максимов Н.А. Методы выделения точек привязки для визуальной навигации беспилотных летательных аппаратов//Труды МАИ. 2012. № 58. С. 6.
5. Порядок ведения Единого государственного реестра недвижимости. Приложение № 1 к Приказу Министерства экономического развития РФ от 16 декабря 2015 г. № 943 «Порядок ведения Единого государственного реестра недвижимости» [Электронный ресурс]. URL: <https://base.garant.ru/71363112/53f89421bbdaf741eb2d1ecc4ddb4c33/>
6. Романенко А.О., Юфряков А.В. Оценка размытия изображения для биометрической идентификации//Наука и образование сегодня. 2018. № 7(30). С. 16-19.

УДК 004.6

МОБИЛЬНОЕ ПРИЛОЖЕНИЕ СИСТЕМЫ АВТОМАТИЗИРОВАННОГО УЧЕТА ПОСЕЩЕНИЯ ЗАНЯТИЙ MOBILE APPLICATION FOR AUTOMATED CLASS ATTENDANCE REGISTRATIONS

Гавриленко С.К., Минасов Ш.М., Мифтахов Б.А.,
ФГБОУ ВО «Уфимский государственный авиационный технический университет»,
г. Уфа, Российская Федерация

S.K. Gavrilenko, Sh.M. Minasov, B.A. Miftakhov,
“Ufa State Aviation Technical University”,
Ufa, Russian Federation

e-mail: job.gavrilenko@yandex.ru

Аннотация. Разработаны модели и алгоритмы процедуры регистрации присутствия студентов на обязательных видах аудиторных занятий для учета выполнения ими графика учебного процесса. Разработанное приложение позволяет

автоматизировать процесс идентификации пользователей, исключая тем самым потерю времени занятий на выполнение переключки. Приложение разработано на платформе Xamarin.Forms в виде мобильного кроссплатформенного приложения и основано на технологии чтения QR-кода, формируемого заранее и размещаемого в презентации, либо генерируемого в режиме реального времени для воспроизведения на любом мобильном устройстве, в т.ч. смартфоне и планшете. При регистрации на занятиях мобильное приложение фиксирует не только факт подтверждения студентом его присутствия, но и фактические геокоординаты местонахождения мобильного устройства, позволяющее исключить некорректное поведение студентов, которые на самом деле не присутствуют на занятии. Сам контроль корректности поведения учитывает формат проведения занятий, допускающий возможность нахождения студента вне аудитории, если мероприятие проводится в дистанционном формате и не подразумевает нахождение всех слушателей в одной аудитории. В этом случае мобильное приложение оценивает факт присутствия на занятии по дополнительным атрибутам, таким как интерактивное участие на занятии, например, в форме краткого опроса мнения, тестирования или рефлексии.

Abstract. Describe models and algorithms for the procedure of registering students' on compulsory auditory lessons to log the presence. The developed application allows automating the process of identification of users, thereby eliminating the loss of class time to per-form roll call. Software solution proposed to implements on Xamarin.Forms platform as mobile cross-platform application, which based on the technology of QR-code reading. The project considers three options: QR-code was generating in real time to show on any mobile device, including smartphone and tablet; integrating early-generated QR-code in presentation and giving printed QR-code on paper. The project propose, that mobile application records not only the fact that the student confirms, but also the actual geo-coordinates of the location of the mobile device, allowing to exclude the incorrect behavior of students who are not actually present in event. The control of correctness of behavior itself takes into account the lessons format and allows for the possibility of finding a student outside the classroom, if the event held in a distance format and does not imply the presence of all students in one auditorium. Proposed algorithm of the mobile application evaluates the fact of presence in the event by additional attributes, such as interactive participation during the lesson, for example, in the form of a brief opinion poll, testing or reflection.

Ключевые слова: образование, учебный процесс, график выполнения учебного процесса, мобильное приложение, кроссплатформенное программное обеспечение, клиент-серверные технологии, системы реального времени.

Keywords: Education, learning process, learning schedule, mobile application, cross-platform software, client-server technologies, real-time systems.

Введение

Важная причина знать поименно всех присутствующих обучающихся на занятиях заключается в применении для оценки уровня освоения компетенций учета посещения ими всех видов аудиторных занятий даже если они проводятся в дистанционном формате, что стало особенно актуальным в период пандемии Covid-19.

В малых группах применяется переключка студентов, но на поточных лекциях такой способ занимает неоправданно много времени, которое следует уделять подаче

материала. Свести потери времени на переключку к нулю с помощью явочных листов решает проблему подачи материала, но:

- 1) обработка явочных листов занимает больше времени, чем переключка,
- 2) обработка производится в отложенном времени и у студентов нет информации о текущем рейтинге,
- 3) в явочных листах часто присутствуют «мертвые души».

Таким образом настоящая работа, посвященная разработке альтернативного метода фиксирования хода учебного процесса, лишённого недостатков вышеперечисленных и иных методов регистрации и идентификации студентов на обязательных к посещению занятиях является актуальной.

Цель исследования – повышение качества учета посещаемости занятий обучающимися за счет снижения затрат времени на его проведение и обеспечения достоверности фиксируемой информации.

Для достижения поставленной цели сформулированы следующие задачи: выполнить поиск информации о существующих отечественных и зарубежных решениях по выше обозначенной проблеме; предложить новый способ решения проблемы учета присутствия обучающихся на занятиях с использованием современных и доступных каждому программно-инструментальных средств; разработать модели и алгоритмы реализации предлагаемого метода; реализовать разработанные алгоритмы в виде программно-аппаратного комплекса и выполнить оценку эффективности его внедрения.

Обзор существующих решений

В результате изучения интернет-источников был найден ряд интересных решений для регистрации пользователей. Среди наиболее популярных и эффективных по мнению интернет-сообщества были рассмотрены системы «Ziplet» [1], «OneTap», «Panorama» [2] и «Student Check in» [3].

В результате анализа указанных решений были выявлены следующие недостатки.

Во-первых, ни одна из этих систем не проверяет геолокацию пользователя в момент сканирования, из-за чего он может легко заявить о своем присутствии в нужной локации, хотя фактически отсутствует в нем.

Во-вторых, три из четырех рассмотренных программ представляли собой Web-приложения, которые плохо функционируют в местах с неустойчивым Интернет-соединением.

В-третьих, только «Ziplet» и «Student Check in» созданы с открытым ядром и являются свободным ПО, в то время как «OneTap» и «Panorama» работают по подписке с ежемесячной оплатой за свои услуги, либо не позволяют пользоваться своим ПО. Также приложения не сохраняют информацию в базе данных, не используют сканирование QR, не совместимы с программным обеспечением университета и т.п.

Результаты сравнительного анализа приведены на рисунке 1.

Название	Работоспособность в России	Проверка геолокации	Сохранение в БД	Офлайн режим	Цена
Ziplet	✓	✗	✗	✗	Бесплатный
Panorama	✗	✗	✓	✗	Доступен только по подписке
OneTap (после недавних событий)	✗	✗	✓	✗	Недоступен для оплаты
Student Check in	✓	✗	✗	✓	Бесплатный

Рисунок 1. Результаты анализа существующих решений

Таким образом готовых решений, удовлетворяющих всем сформулированным требованиям, не оказалось.

Моделирование системы

Как показал анализ открытых источников по обозначенной проблеме, системы, которая бы могла удовлетворить наши потребности в полном объеме не оказалось. Таким образом был сделан вывод о том, что ее придется разрабатывать самостоятельно.

Исходя из настоящих реалий сделано предположение, что систему регистрации явки студентов на занятия с гарантией отсутствия возможности регистрации физически отсутствующих лиц можно сделать на основе персонально выдаваемых идентификаторов со случайно-генерируемым кодом.

Такая технология применялась банками в виде генерируемых конкретному клиенту пула одноразовых паролей. В нашем случае такой пароль необходимо генерировать для неопределенного пользователя. Идентификация пользователя состоится по регистрируемым в мобильном приложении данным, а факт физического присутствия на месте проведения мероприятия подтверждается сканированием одноразового QR-кода.

На рисунке 2 показан фрагмент функциональной диаграммы работы системы. Кроме персонального идентификатора мобильное приложение передает на сервер геокоординаты пользователя и при разнице с геопозицией лектора система определяет корректность полученной информации. Кодирование ключевой фразы для регистрации явки в виде QR-кода позволяет повысить защиту от некорректно-вносимых данных и скорость ввода информации, поскольку набор пароля на экране смартфона – процедура не из элементарных.

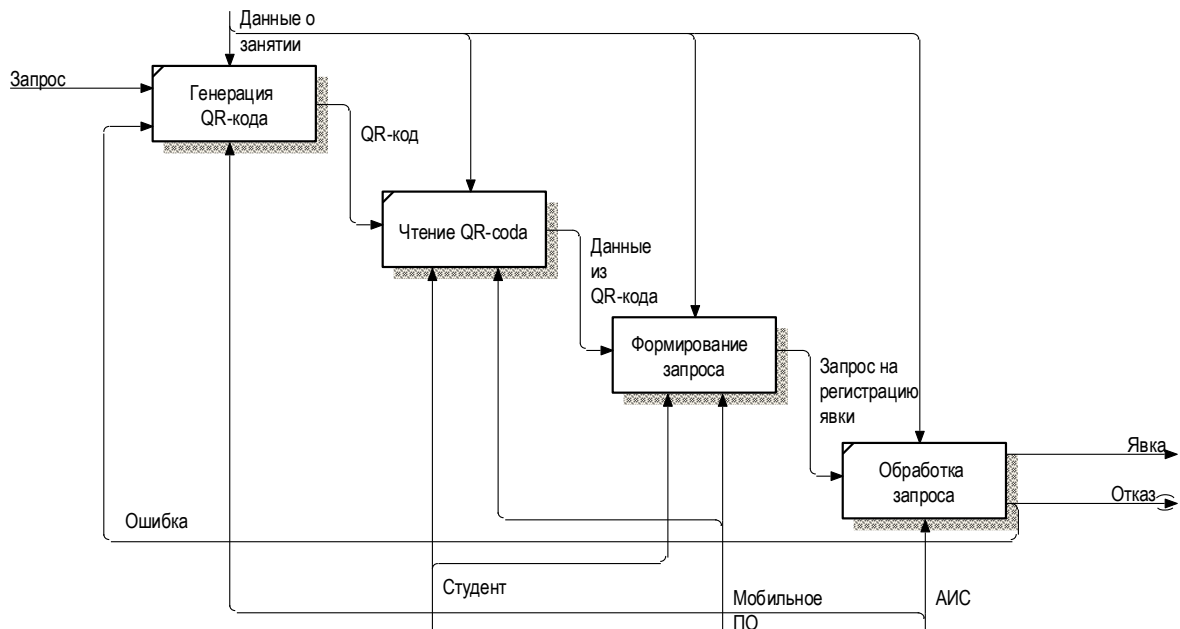


Рисунок 2. Фрагмент функциональной модели системы

Алгоритм работы системы

Укрупненный алгоритм работы серверной части системы представлен на рисунке 3.

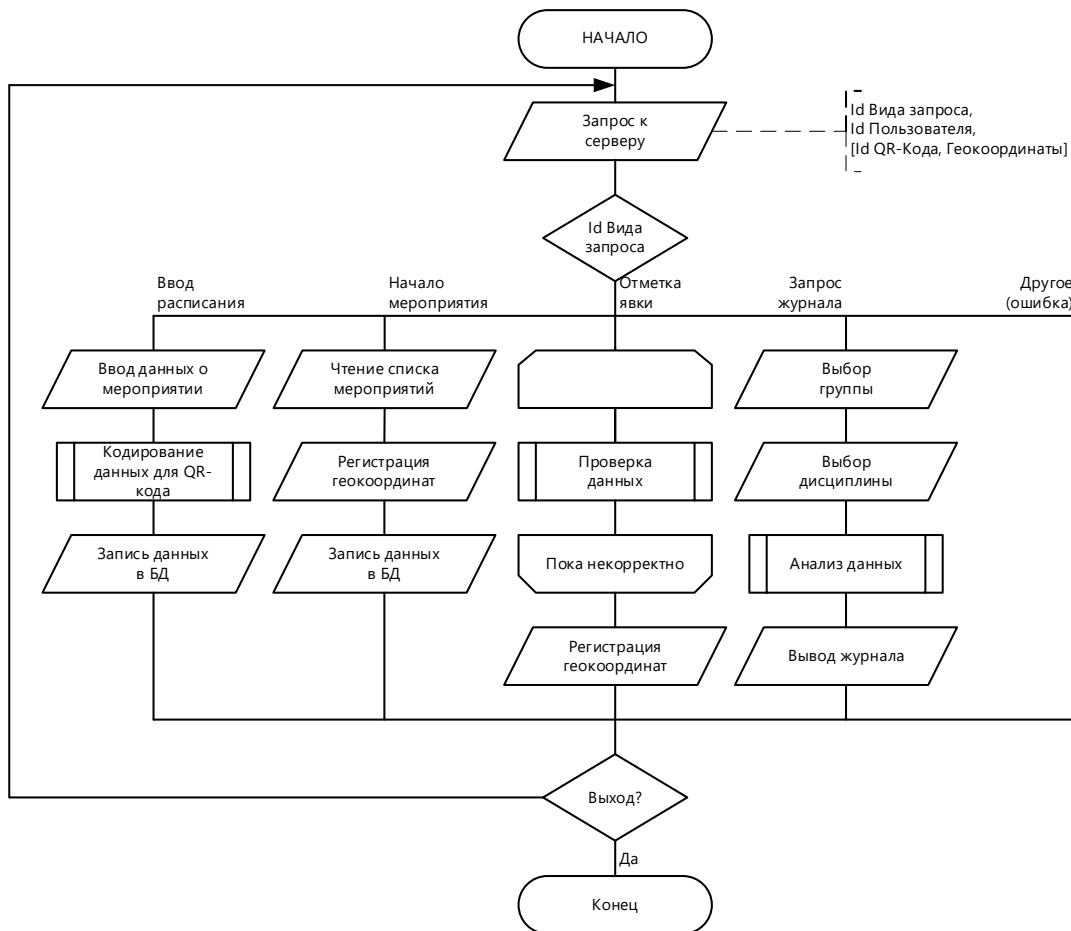


Рисунок 3. Укрупненный алгоритм работы серверной части системы

Первоначально преподаватель взаимодействует с сервером приложений через мобильное или веб-приложение, и вводит данные о проводимых занятиях. В начале занятия преподаватель с мобильного устройства активирует систему регистрации, при этом в систему поступает информация о его геопозиции, получаемой с мобильного устройства. На мобильное устройство преподавателя (удобнее использовать планшетный ПК) выводится QR-код с зашифрованными в нем данными таким образом, чтобы было невозможно выделить из них информацию о геопозиции мероприятия, что исключает подделку запросов на обход требований очного присутствия. Если занятие проводится в аудитории, то QR-код не может быть повторно использован и для следующего студента генерируется новый.

Мобильное приложение студента сканирует QR-код и отправляет запрос на сервер, дополняя полученные сведения информацией о своей геопозиции. В случае несоответствия геокоординат сохраненных в QR-коде координатам фактического нахождения студента, факт его присутствия не подтверждается. При этом для распределенных занятий и занятий в дистанционной форме такое несоответствие игнорируется.

Архитектура системы

Система построена по трехзвенной архитектуре. Мобильное приложение – серверное приложение – база данных. Архитектура разработанной системы учета посещения занятий обучающимися представлена на рисунке 4.

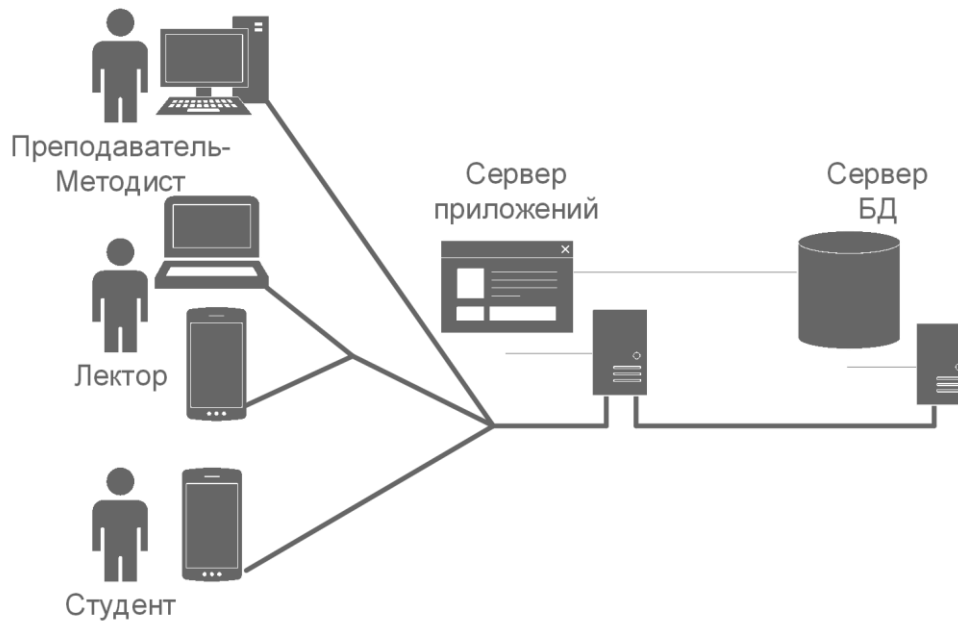


Рисунок 4. Архитектура системы

Программные инструменты реализации проекта

Для реализации мобильного приложения были рассмотрены несколько фреймворков, по результатам анализа которых выбор остановился на: ZXing.Mobile, Xamarin.Forms.Xaml, Xamarin.Essentials, System.Net.Http, Newtonsoft.Json.

Серверная часть системы реализована на стеке приложений Apache, Php, Maria DB. Для работы с QR-кодом использована php-библиотека, распространяемая в открытых исходных кодах.

После проведения тестирования работоспособности и устранения ряда системных ошибок было получено полностью автономное кроссплатформенное приложение, позволяющее эффективно проводить контроль явки студентов на занятия.

Оценка работоспособности предложенного решения

Апробация прототипа системы была произведена на научной конференции «Неделя науки» и ряде учебных занятий. Погрешность определения геокоординат лишь в одном случае превысила 10 м, но причиной этого было отключенное из-за низкого заряда батареи геопозиционирование.

Для гарантии корректной работы приложения был оценен пиковый трафик, который может поступать на сервер.

По выполненной оценке, максимальный объем трафика для нашего университета теоретически может составить 10000 запросов по 16 байт одномоментно, что составит всего 1,2 Мбита.

Таким образом приложение благодаря его реализации в качестве мобильного не будет оказывать недопустимо высокой нагрузки на сервер и каналы связи.

Оценка экономической эффективности внедрения

По примерной оценке, на основании средних показателей по аудиторной нагрузке на одного преподавателя в 720 часов в год, при двухчасовых занятиях в группах и 4-х часовых в подгруппах получается, что преподаватель за год должен обработать 360 таких

документов. При средних затратах по 6 минут на страницу текста суммарное время составит около 36 часов или одну рабочую неделю.

При выполнении требований законодательства и майских указов В.В. Путина 2012 г. [4] это время должно оплачиваться в Башкирии суммой порядка 20 000 рублей каждому преподавателю, что в рамках среднестатистического университета с 800 преподавателями складывается в 16 миллионов рублей.

По данным [5] на середину 2021 года количество преподавателей в Российской Федерации составило свыше 300 000 человек. Оценочно сумма затрат на оплату бессмысленно потраченного времени составит порядка шести миллиардов рублей в год.

Что касается затрат на разработку, то на создание прототипа потребовалось порядка 12 часов времени работы программиста. При стоимости труда программиста в 120 000 рублей в месяц, затраты на разработку и отладку финальной версии со всеми налогами не превысит этой суммы.

Эксплуатационные затраты от внедрения этой системы ничтожны как в аппаратном, так и в программном смысле, поскольку все необходимые программно-аппаратные решения в вузах давно развернуты, а само приложение не создает большой нагрузки.

Выводы

В результате исследования было показано, что цифровизация образовательного процесса даже в мелочах может существенно сократить непроизводительные затраты и повысить эффективность работы специалистов, привлекаемых к образовательному процессу.

В рамках настоящего проекта рассматривается его развитие на предмет расширения функционала, который позволит повысить качество учебного процесса за счет устранения большего числа коллизий, приводящих к напрасной потере времени. С учетом снижения доли контактной работы в пользу СРС ценность этого времени растет, и каждая минута становится драгоценной, ведь несмотря на то, что в длинных дисциплинах с аудиторной нагрузкой 36-72 часа потеря одного часа не так существенна, то в дисциплинах, в которых время лекционной работы сократилось до 6 часов, это катастрофа.

Литература

1. Check in with your students. Zipler official Web-site. URL: <https://zipler.com> (Дата обращения: 02.04.2022).
2. Student Check In. Official Web-site. URL: <https://www.studentcheckin.com/> (Дата обращения: 02.04.2022).
3. Panorama Education. Official Web-site. URL: <https://www.panoramaed.com> (Дата обращения: 02.04.2022).
4. Указ Президента Российской Федерации от 07.05.2012 г. № 597 О мероприятиях по реализации государственной социальной политики.
5. Знания в приоритете. Российская газета. Спецвыпуск от 15.06.2021. URL: <https://rg.ru/gazeta/rg-spec/2021/06/15/5.html> (Дата обращения: 02.04.2022).

УДК 004.681.3

**ИСПОЛЬЗОВАНИЕ ЯЗЫКА ЗАПРОСОВ SQL
ПРИ ПРОЕКТИРОВАНИИ ИНФОРМАЦИОННЫХ СИСТЕМ****USING THE SQL QUERY LANGUAGE
WHEN DESIGNING INFORMATION SYSTEMS**

¹Каданцев М.Н., ¹Султанова Е.А., ²Филиппова Е.В.,
¹Уфимский государственный нефтяной технический университет,
ул. Космонавтов, 1, г. Уфа, Республика Башкортостан, 450064, Россия
²МБОУ «Гимназия №82», г. Уфа, Российская Федерация

M.N. Kadantsev¹, E.A. Sultanova¹, E.V. Filippova²,
¹Ufa State Petroleum Technological University,
Kosmonavtov Str., 1, Ufa, Republic of Bashkortostan, 450064, Russia
²MBEI "Gymnasium No. 82", Ufa, Russian Federation

e-mail: mail_im3435@mail.ru

Аннотация. Как одно из приоритетных направлений на текущий момент разработка и эксплуатация банков и баз данных в информационных системах, а также разработка приложений на основе систем управления базами данных (СУБД) являются основой практической реализации, отражающей современное представление о процессах преобразования и потребления информации в информационном обществе. Но не во всех прикладных областях данное направление нашло полное своё применение, в частности в нефтегазовом деле. В данной статье рассмотрены вопросы в сфере реализации информационных систем, связанные с применением информационных технологий в создании базы данных с использованием языка структурированных запросов SQL, причем акцент сделан на разработке информационной системы для автоматизации анализа и учета, данных расхода горюче-смазочных материалов (ГСМ). Язык SQL позволяет принимать SQL-запросы от других компонентов СУБД и пользовательских приложений. SQL – мощный инструмент, который обеспечивает пользователям, программам и вычислительным системам доступ к информации, содержащейся в реляционных базах данных. Вместе с тем надо отметить ряд недостатков присущие языку SQL: несоответствие реляционной модели данных, сложность, отступления от стандартов.

Abstract. As one of the priority directions at the moment, the development and operation of banks and databases in information systems, as well as the development of applications based on database management systems (DBMS) are the basis of practical implementation that reflects the current understanding of the processes of transformation and consumption of information in the information society. However, this direction has not found its full application in all applied areas, in particular in the oil and gas industry. This article discusses issues in the field of information systems implementation related to the use of information technologies in creating a database using the structured query language SQL, with emphasis on the development of an information system for automating the analysis and accounting of fuel and lubricants consumption data. The SQL language allows you to accept SQL queries from other DBMS components and user applications. SQL is a powerful tool that provides users, programs, and computing systems with access to information contained in

relational databases. At the same time, we should note a number of disadvantages inherent in the SQL language: inconsistency of the relational data model, complexity, and deviations from standards.

Ключевые слова: информационные системы, база данных, SQL, горюче-смазочные материалы.

Keywords: information systems, database, SQL, fuel and lubricants.

Обоснованный выбор моделей для проектирования информационной системы (ИС) является одной из основных задач этапа концептуального проектирования (информационной структуры программного обеспечения ИС). Можно использовать различные модели, в частности *модель «сущность – связь»* [1].

Из моделей типа «сущность – связь» наиболее известна модель П. Чена, или ER–модель (Entity-Relationship). На языке ER - модели концептуальная схема может быть представлена ERD (ER-*диаграммой*), в которой множество сущностей обозначается прямоугольниками, множество связей – ромбами.

Как правило, при использовании систем управления базами данных в информационных системах, после этапа моделирования данных выполняется преобразование ER–модели в реляционную модель, т.е. сущностям и связям ER-модели ставятся в соответствие отношения, или таблицы, которые входят в состав базы данных.

Для выполнения запросов применяется язык структурированных запросов SQL, который является средством извлечения из реляционной базы данных информации, необходимой пользователю. В данной статье рассматриваются возможности SQL по выборке данных из базы данных.

Вся информация, описывающая конкретную предметную область, должна быть определенным образом абстрагирована и формализована. Основными направлениями формализации информации о предметной области являются:

- теория классификации, базирующаяся на таксономическом и мерономическом описании информации. Таксономическое описание основано на идеологии множеств, а мерономическое осуществляется через строго формализованное определение классов;
- теория измерений, предлагающая базу для качественных и количественных измерений через классификационные и порядковые шкалы;
- семиотика, изучающая знаковые системы с точки зрения синтактики, семантики и прагматики.

Информационная система состоит из системы управления БД, системы управления базой моделей (СУБМ) и системы управления интерфейсом между пользователем и компьютером (рис. 1).



Рисунок 1 Информационная система

Процесс обработки данных играет важную роль, так как информация, введенная в хранилище данных, может использоваться в дальнейшем для принятия решений и предотвратить повторный ввод некорректных данных. Диагностика основных фондов требует устойчивого и комплексного подхода к формированию системы.

Исходя из методик, составляется, возможная блок-схема с алгоритмом работы (рисунок 2), где на выходе выводится отчет об оценке текущего состояния оборудования или сооружения. Для обработки информации используются материалы, подготовленные и представленные в базе данных.

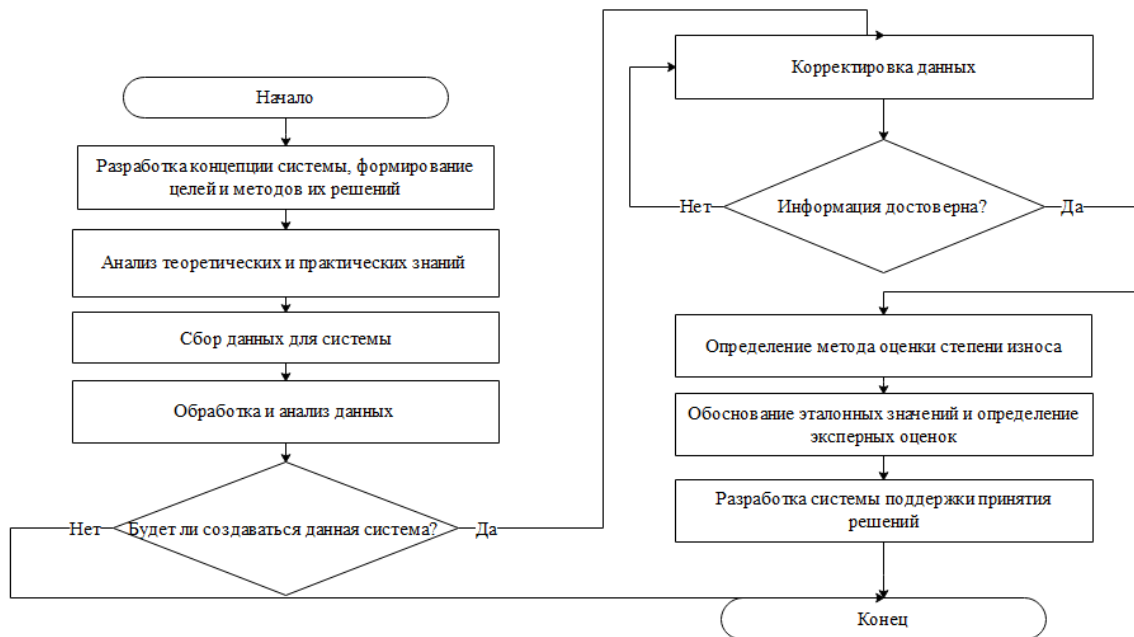


Рисунок 2. Примерная блок-схема алгоритма работы системы

Информация о возможных ошибках и о вводе корректных данных записываются в отдельные базы данных, для наглядного вывода информации в виде записи данных, а также для использования материалов в дальнейшем. Система осуществляет контроль отклонений различных параметров и выдает отчеты, позволяющие в явной форме выводить информацию о различных решениях, принятых системой. Результаты анализа передаются в модуль сопоставления данных, где происходит поиск оптимальных решений, определение приоритетов в замене оборудования и дальнейшее планирование работ с оборудованием и изделиями.

При проектировании информационной системы необходимо из описания предметной области извлечь информацию и выделить сущности. Выполнить анализ предметной области, т.е. определить объекты предметной области и связи между объектами. При выборе состава и структуры предметной области возможны два подхода: функциональный и предметный.

Функциональный подход реализует принцип движения «от задач» и применяется, когда определен комплекс задач, для обслуживания которых создается информационная система. В этом случае можно выделить минимальный необходимый набор объектов предметной области, которые должны быть описаны.

В предметном подходе объекты предметной области определяются с таким расчетом, чтобы их можно было использовать при решении множества разнообразных,

заранее не определенных задач. Чаще всего используется комбинация этих двух подходов.

Опишем предметную область, предлагаемую в качестве примера в проекте, реализация которого представлена в данной статье.

Предположим, что для бухгалтерии требуется разработать информационную систему для автоматизации анализа и учета, данных расхода горюче-смазочных материалов (ГСМ) [2].

Учет расходов ГСМ традиционно является серьезной задачей для большинства практикующих бухгалтеров, так как он обусловлен необходимостью контроля за расходом топлива, который связан с большим количеством трудно учитываемых параметров. К ГСМ относят следующие группы:

- все виды топлива (бензин, дизельное топливо, природный и нефтяной газ);
- смазочные вещества (моторные, трансмиссионные и специальные масла, пластичные смазки);
- используемые при уходе за автотранспортом и в его работе жидкости (тормозная, охлаждающая, тосол и антифриз).

Предельных норм для отнесения ГСМ на расходы не существует. Статья 25 НК РФ требует только документального подтверждения фактически произведенных расходов на ГСМ. Но при этом организация обязана издать собственный внутренний документ, устанавливающий предельно допустимые нормы расходов с учетом эксплуатационных характеристик техники. Такие нормы должны быть установлены для техники в период ее эксплуатации, ремонта и техобслуживания.

Достаточно сложно самостоятельно рассчитать такие нормы, необходимо измерять расход ГСМ в пробках, при разных режимах эксплуатации, в разную погоду. Не все технические службы предприятий могут это сделать и предоставить данные, которые и лягут в основу разработки учетной политики организации, связанной с ГСМ.

Именно поэтому и возникает необходимость в программных продуктах, уже содержащих разработанные и проверенные нормы расходования ГСМ. Программа для расчета расхода топлива должна исключить необходимость ввода этих показателей вручную [3-5].

Программа, учитывающая ГСМ, обычно должна иметь возможность сформировать и распечатать путевые листы, а также обработать содержащиеся в них данные для расчета норм ГСМ.

В настоящее время используются стандартные системы учета ГСМ:

- в системах 1С, например, «1С: Бухгалтерия 7.7». В рамках этой знакомой программы разработана отдельная конфигурация, позволяющая учитывать затраты на ГСМ. Она помогает рассчитать расходы на топливо, но, нормы расходования необходимо брать из методических рекомендаций или учетной политики и вводить в программу в ручном режиме. Модуль позволяет вести не только бухгалтерский учет ГСМ, но и оперативный, используемый для ежедневного планирования, и позволяет формировать различные виды отчетов;

- для государственных учреждений в рамках 1С разработан модуль: «Автотранспорт и учет ГСМ для государственных учреждений». Он предназначен для бюджетных организаций, но при определенной доработке может быть интегрирован и в стандартную базу. Эта программа для учета расхода топлива подойдет для небольшого автопарка, но не поможет работе автотранспортных предприятий. Для них разрабатываются свои продукты;

- существуют и отдельные программы, позволяющие рассчитать расходы по ГСМ и произвести учет путевых листов. Достаточно часто они способны интегрироваться с

ИС или в них предусмотрен механизм передачи данных в этот программный комплекс. Такие программы, как правило, применяются компаниями с большим автопарком.

В этой статье на примере конкретной разработки будут рассмотрены идеи и приемы, используемые при создании информационной системы по учету расхода ГСМ. Данная ИС имеет может иметь практическое применение при условии дальнейшей доработки.

Основным объектом *предметной области* является бухгалтерия (учет расхода ГСМ). Мы будем придерживаться функционального подхода, т.е. идти от задач, которые будут решаться.

Основные предметно-значимые сущности: горюче-смазочные материалы, подразделения.

Основные предметно-значимые атрибуты сущностей:

- горюче-смазочные материалы – название;
- подразделения – название.

Основные требования к функциям системы:

- вывести расходы на горюче-смазочные материалы по подразделениям;
- вывести расходы на горюче-смазочные материалы по видам ГСМ.

Выполним решение поставленной задачи в системе управления базами данных Microsoft Access в программном продукте Microsoft Office с использованием языка запросов SQL [6].

Выводы

Проанализированы существующие на данный момент стандартные системы учета ГСМ, с помощью которых можно обратиться и ознакомиться с методиками учета ГСМ, на основании чего выявлена и обоснована необходимость применения информационных систем и разработок в данной области. В статье авторы предлагают один из подходов решения задач, возникающих в процессе проектирования информационной системы, который позволит упростить, структурировать, автоматизировать работу с использованием языка структурированных запросов SQL по созданию базы данных и обозначить наиболее существенные задачи. В качестве такой задачи рассматривается автоматизация анализа и учета, данных расхода горюче-смазочных материалов.

Литература

1. Введение в информационные системы производства [Электронный ресурс]: учебно-методическое пособие к практическим занятиям по дисциплине «Введение в информационные системы производства»/УГНТУ, Салават. филиал., каф. ОНД; сост. Т. М. Левина. – Салават: УГНТУ, 2017.
2. Задания к выполнению лабораторных работ по курсу «Информатика»: базы данных. Для студентов всех специальностей дневной формы обучения/Составитель: М.В. Додонов. – Самара: СамГАПС, 2004. – 16 с.
3. Vulfin A.M., Frid A.I., Giniyatullin V.M. Neuralbase model for detection and recognition of technological situations within the scope of data mining strategy//Optical Memory & Neural Networks (Information Optics). 2010. Т. 19. № 3. С. 207-212.
4. Гиниятуллин В.М., Салихова М.А., Чурилов Д.А. Троичная логика. Базис, совершенная форма, правила минимизации//Перспективы науки. 2019. № 12 (123). С. 28-32.

5. Гиниятуллин В.М., Салихова М.А. Эффект компенсации ошибок округления в троично-сбалансированной системе счисления//Вестник кибернетики. 2020. № 4 (40). С. 14-20.

6. Язык запросов Transact SQL [Электронный ресурс]: учебно-методическое пособие к выполнению практических и лабораторных работ для студентов направления 230100 «Информатика и вычислительная техника». УГНТУ, каф. ВТИК; сост.: О.П. Тулупова, В.Р. Ганиева. – Уфа: УГНТУ, 2014. – 440 Кб.

УДК 004.94

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПЕЧИ В СРЕДЕ SIMINTECH

MATHEMATICAL MODEL OF THE FURNACE IN SIMINTECH

Кирюшин О.В., Мухамедьяров А.А.,
Уфимский государственный нефтяной технический университет,
ул. Космонавтов, 1, г. Уфа, Республика Башкортостан, 450064, Россия

O.V. Kiryushin, A.A. Muhamediarov,
Ufa State Petroleum Technological University,
Kosmonavtov Str., 1, Ufa, Republic of Bashkortostan, 450064, Russia

e-mail: ainour.mm@mail.ru

Аннотация. Трубчатая печь – это сложный тепловой объект, который предназначен для нагрева сырья. Такие печи пользуются большим спросом в нефтегазоперерабатывающей, химической и нефтехимической промышленности. По конструктивному признаку трубчатые печи делятся на радиантные, конвекционные и радиантно-конвекционные. В радиантных печах все трубы по пути следования сырья, размещены в камере сгорания, таким образом основной нагрев происходит путем радиации. Такой тип печей используется при нагреве небольшого количества сырья до температуры 300 °С. В конвекционных печах тепло передается сырью конвекцией, что обеспечивается путем омывания труб змеевика дымовыми газами. Такие печи применяются для нагрева чувствительных к высоким температурам сред, требующих мягких условий нагрева. В радиантно-конвекционных трубчатых печах большая часть всего тепла (40-60 %) передается радиацией, а остальная – конвекцией. Сложность трубчатой печи заключается в значительных объемах, по которым распределяются температурные поля, которые нужно учитывать при управлении в целях недопущения закоксовывания и прогара. В случае, когда провести опыты на реальном объекте не предоставляется возможным из-за технических или экономических соображений, используются математические модели, с помощью которых моделируется поведение данных объектов с целью разработки и оптимизации алгоритмов управления. В данной работе рассматривается математическая модель печи и ее реализация в виде программного имитатора в среде SimInTech.

Abstract. A tubular furnace is a complex thermal object that is designed to heat raw materials. Such furnaces are in great demand in the oil and gas processing, chemical and petrochemical industries. By design, tubular furnaces are divided into radiant, convection and radiant-convection. In radiant furnaces, all pipes along the route of the raw materials are placed

in the combustion chamber, so the main heating occurs by radiation. This type of furnace is used when heating a small amount of raw materials to a temperature of 300 °C. In convection ovens, heat is transferred to raw materials by convection, which is provided by washing the pipes of the coil with flue gases. Such furnaces are used for heating environments sensitive to high temperatures that require mild heating conditions. In radiant-convection tubular furnaces, most of the total heat (40-60%) is transmitted by radiation, and the rest by convection.

The complexity of a tubular furnace lies in the significant volumes over which temperature fields are distributed, which must be taken into account when controlling in order to prevent coking and burnout. In the case when it is not possible to conduct experiments on a real object due to technical or economic considerations, mathematical models are used to model the behavior of these objects in order to develop and optimize control algorithms.

In this paper, we consider a mathematical model of the furnace and its implementation in the form of a software simulator in the SimInTech environment.

Ключевые слова: моделирования, трубчатая печь, Simintech, имитатор, передаточные функции.

Keywords: modeling, tube furnace, Simintech, simulator, transfer functions.

Прежде, чем реализовывать математическую модель в среде SimInTech, необходимо в соответствии с идеями системного подхода при моделировании составить концептуальную (рисунок 1) и топологическую (рисунок 2) модели печи [1].

Модель печи далее представляется как система инерционных звеньев первого порядка с запаздыванием:

$$W = K \frac{1}{Ts+1} e^{-\tau s}, \quad (1)$$

где K – коэффициент усиления модели;

T – постоянная времени инерционного звена;

τ – «чистое» запаздывание модели.

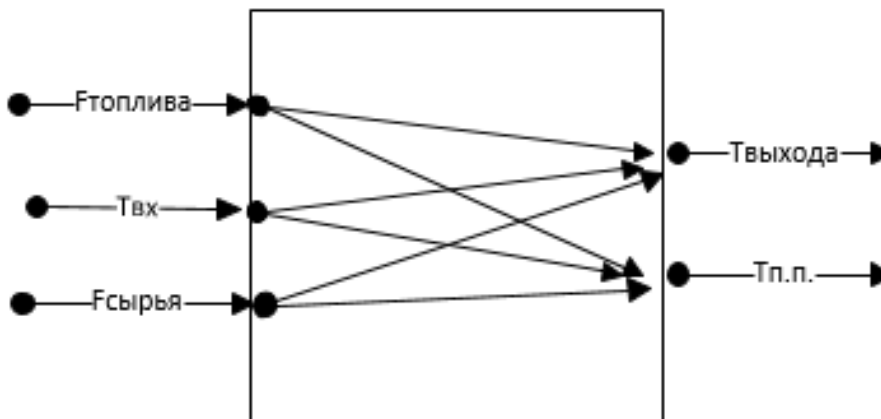


Рисунок 1. Концептуальная модель печи

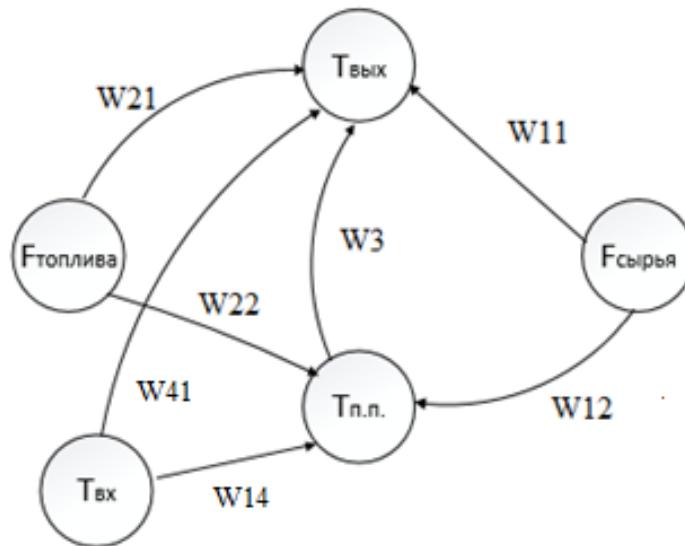


Рисунок 2. Топологическая модель печи

Постоянные времени и «чистое» запаздывание были определены эмпирически. Коэффициенты усиления были подобраны с учетом приоритета влияния на изменение выходных величин.

Для определения коэффициентов усиления записываются уравнения статики для выходных величин:

$$T_{\text{вых}} = -K_{11} \cdot F_{\text{сырья}} + K_3 \cdot T_{\text{п.п.}} + K_{21} \cdot F_{\text{топлива}} + K_{41} \cdot T_{\text{вх}}, \quad (2)$$

$$T_{\text{п.п.}} = -K_{12} \cdot F_{\text{сырья}} + K_{22} \cdot F_{\text{топлива}} + K_{14} \cdot T_{\text{вх}}, \quad (3)$$

где $F_{\text{сырья}}$ – номинальное значение расхода сырья в печь, м³/ч;

$T_{\text{п.п.}}$ – номинальное значение температуры на перевале, °С;

$F_{\text{топлива}}$ – номинальное значение расхода топлива на горелки, м³/ч;

$T_{\text{вх}}$ – номинальная температура сырья на входе в печь, °С;

$T_{\text{вых}}$ – номинальная температура на выходе из печи, °С.

Коэффициенты K_{11} и K_{12} берутся с минусами, так как, увеличивая расход сырья, температура на выходе уменьшается.

Предположим, что на выходную температуру из печи $T_{\text{вых}}$ влияют [2] следующие параметры: расход сырья $F_{\text{сырья}}$ на минус 15 %, температура на перевале $T_{\text{п.п.}}$ на 15 %, расход топлива $F_{\text{топлива}}$ на 30 %, температура на входе печи на $T_{\text{вх}}$ 70 %.

Предположим, что на температуру на перевале $T_{\text{п.п.}}$ влияют параметры: расход сырья $F_{\text{сырья}}$ на минус 10 %, расход топлива $F_{\text{топлива}}$ на 60 %, температура на входе печи $T_{\text{вх}}$ на 50 %.

Тогда коэффициенты усиления K_{11} , K_{12} , K_{21} , K_{22} , K_3 определяются по формуле:

$$K = \frac{a \cdot b}{c \cdot 100 \%}, \quad (4)$$

где a – номинальное значение выходного параметра;

b – влияние входного параметр на выходной, %;

c – номинальное значение входного параметра.

Номинальные значения параметров:

$$F_{\text{сырья}} = 80 \text{ м}^3/\text{ч}, \quad F_{\text{топлива}} = 2000 \text{ м}^3/\text{ч}, \quad T_{\text{п.п.}} = 750 \text{ }^\circ\text{C},$$

$$T_{\text{вх.}} = 365 \text{ }^\circ\text{C}, \quad T_{\text{вых.}} = 475 \text{ }^\circ\text{C}.$$

Коэффициенты усиления согласно формуле (4) имеют вид:

$$K_{11} = \frac{T_{\text{вых.}} \cdot 20 \%}{F_{\text{сырья}} \cdot 100 \%} = \frac{475 \cdot 15 \%}{80 \cdot 100 \%} = 0,89 \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$K_{21} = \frac{T_{\text{вых.}} \cdot 80 \%}{F_{\text{топлива}} \cdot 100 \%} = \frac{475 \cdot 30 \%}{2000 \cdot 100 \%} = 0,071 \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$K_{12} = \frac{T_{\text{п.п.}} \cdot 30 \%}{F_{\text{сырья}} \cdot 100 \%} = \frac{750 \cdot 10 \%}{80 \cdot 100 \%} = 0,9375 \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$K_{22} = \frac{T_{\text{п.п.}} \cdot 130 \%}{F_{\text{топлива}} \cdot 100 \%} = \frac{750 \cdot 60 \%}{2000 \cdot 100 \%} = 0,225 \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$K_3 = \frac{T_{\text{вых.}} \cdot 40 \%}{T_{\text{п.п.}} \cdot 100 \%} = \frac{475 \cdot 15 \%}{750 \cdot 100 \%} = 0,095.$$

Будем считать, что температура на входе $T_{\text{вх}}$ – это постоянная величина, поэтому $K_{41} = K_{14} = 1$. Тогда передаточные функции согласно формулам (1) и (4) имеют вид:

$$W_{11} = \frac{0,89}{4s+1} e^{-3s} \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$W_{12} = \frac{0,9375}{2s+1} e^{-2s} \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$W_{21} = \frac{0,071}{3s+1} e^{-3s} \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$W_{22} = \frac{0,225}{2,5s+1} e^{-2s} \frac{^\circ\text{C}}{\text{м}^3/\text{ч}},$$

$$W_3 = \frac{0,095}{1,8s+1} e^{-2s},$$

$$W_{41} = 1,$$

$$W_{14} = 1.$$

Передаточная функция W_{11} характеризует влияние расхода сырья в печь на температуру на выходе из печи, W_{12} – влияние расхода сырья в печь на температуру на перевале печи, W_{21} – влияние расхода топлива на горелках на температуру на выходе из печи, W_{22} – влияние расхода топлива на горелках на температуру на перевале печи, W_3 – влияние температуры на перевале печи на температуру на выходе из печи, W_{41} – влияние начальной температуры сырья печи на температуру на выходе из печи, W_{14} – влияние начальной температуры сырья печи на температуру на перевале печи.

По полученным передаточным функциям далее составляется имитационная модель в среде SimInTech [3]. Полученная модель представлена на рисунке 3.

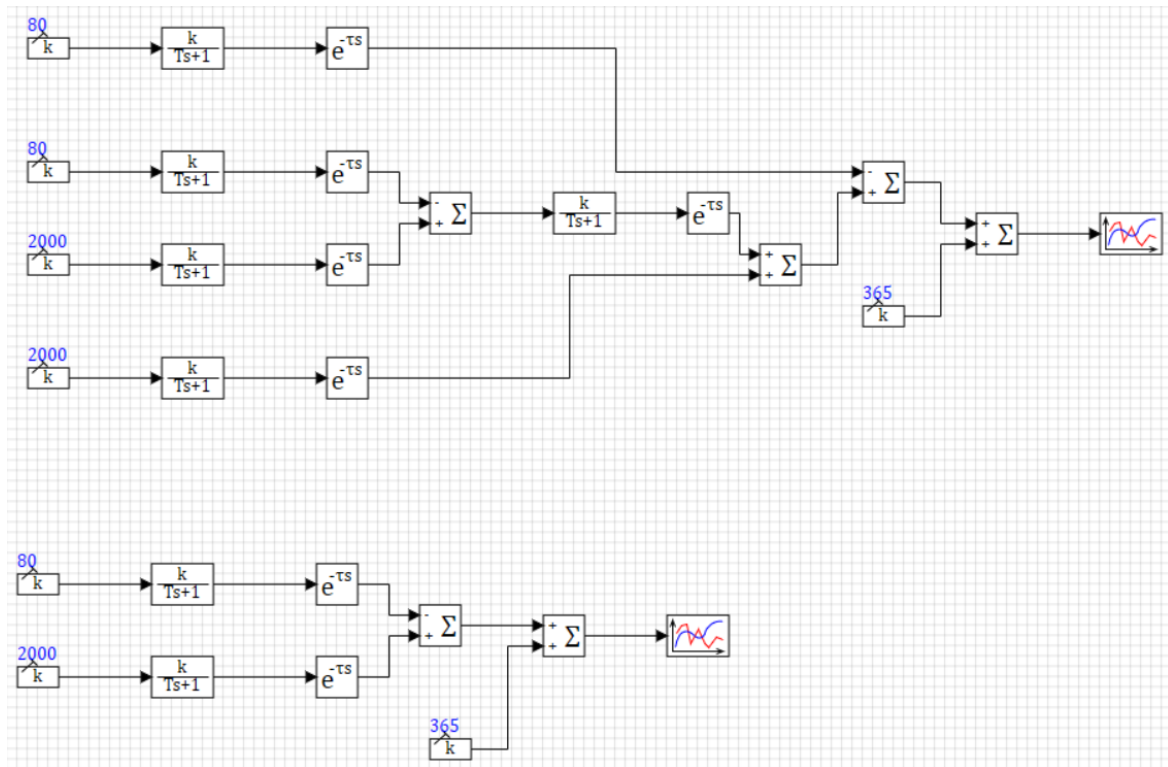


Рисунок 3. Имитационная модель печи

Для составления имитационной модели печи в среде SimInTech были использованы следующие блоки: «Константа», «Сумматор», «Инерционное звено 1-го порядка», «Идеальное транспортное запаздывание», «Временной график».

Блок «Константа» формирует на выходе заданную пользователем постоянную величину $y(t) = k$. С помощью данного блока были заданы номинальные значения параметров.

Блок «Временной график» реализует отображение текущих результатов моделирования в виде временных зависимостей для одной или одновременно для нескольких переменных.

Блок «Сумматор» реализует операцию алгебраического поэлементного суммирования входных сигналов с учетом весовых коэффициентов, которые задаются в свойствах блока.

Блок «Инерционное звено 1-го порядка» реализует математическую модель звена, динамика которого описывается линейным обыкновенным дифференциальным уравнением (ОДУ) вида

$$T_i \frac{dy_i(t)}{dt} + y_i(t) = k_i x_i(t).$$

Коэффициент усиления k_i и постоянная времени T_i задаются в свойствах блока (рисунок 4) в столбце «Значение» в соответствующих строках.

Свойства : Aperiodika6

Свойства	Параметры	Общие	Порты	Визуальные слои
Название	Имя	Формула	Значение ▲	
Коэффициенты усиления	k		[0.89]	
Начальные условия	x0		[0]	
Постоянные времени	T		[4]	

Рисунок 4. Свойства блока «Инерционное звено 1-го порядка»

Блок «Идеальное транспортное запаздывание» реализует математическую модель звена, динамика которого описывается линейным уравнением с запаздыванием:

$$y_i(t) = x_i(t - T_i).$$

Время запаздывания T_i задается в свойствах блока (рисунок 5) в столбце «Значение».

Свойства : TimeDelay9

Свойства	Общие	Порты	Визуальные слои
Название	Имя	Формула	Значение ▲
Максимальный размер буфера	maxsize		0
Начальный размер буфера	stacksize		4096
Шаг дискретизации задержки, сек	distau		[0]
Время запаздывания	tau		[3]
Метод интерполяции	interpmethod		Линейная
Фиксированный размер буфера	fixedbuffer		<input type="checkbox"/> Нет
Способ инициализации	stackinit		По входу

Рисунок 5. Свойства блока «Идеальное транспортное запаздывание»

Результатом имитационного моделирования стали переходные кривые процесса, построенные в среде SimInTech (рисунки 6 и 7). Графики достаточно адекватно отображают поведение реального объекта, что дает возможность далее проводить имитационные эксперименты, например, направленные на разработку и оптимизацию алгоритмов управления технологическим процессом.

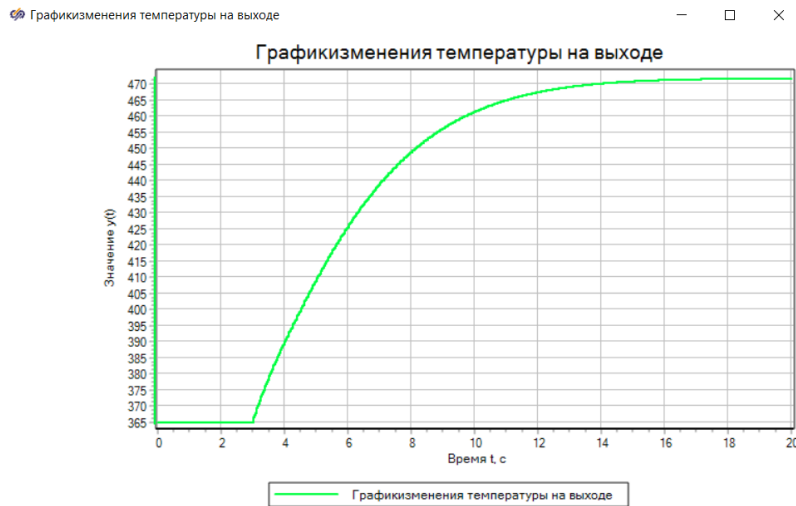


Рисунок 6. График изменения температуры на выходе

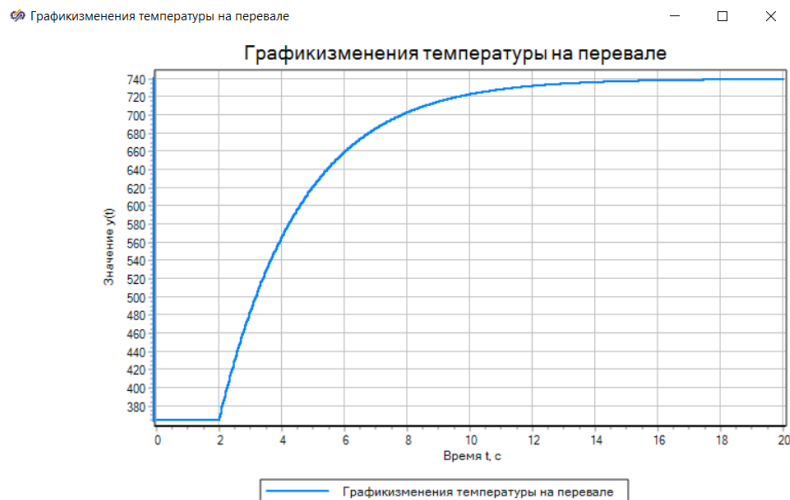


Рисунок 7. График изменения температуры на перевале

Выводы

SimInTech – среда разработки математических моделей, алгоритмов управления, интерфейсов управления и автоматической генерации кода для контроллеров управления и графических дисплеев. Среда обладает всеми возможностями для разработки и реализации имитационных моделей технологических процессов, на основе которых можно производить различные исследования, например, в области оптимизации алгоритмов управления.

Литература

- 1 Вережкин, А.П. Автоматизация технологических процессов и производств в нефтепереработке и нефтехимии/А.П. Вережкин, О.В. Кирюшин. – Уфа. Изд-во УГНТУ, 2005 – 171 с.
- 2 Ахметов, С.А. Лекции по технологии глубокой переработки нефти в моторные топлива: Учебное пособие. – СПб.: Недра, 2007. – 312 с.
- 3 Хабаров, С.П. Основы моделирования технических систем. Среда Simintech/С.П. Хабаров, М.Л. Шилкина. – СПб.: Лань, 2019. – 120 с.

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ,
УПРАВЛЕНИИ И БИЗНЕСЕ**

УДК 004.912

**ПОДДЕРЖКА ПРОЦЕССОВ ИНФОРМАЦИОННОГО УПРАВЛЕНИЯ
С ПОМОЩЬЮ ПРОГРАММНОГО СЕРВИСА
ДЛЯ АВТОМАТИЗИРОВАННОЙ ОБРАБОТКИ
СЛАБОСТРУКТУРИРОВАННЫХ ТЕКСТОВЫХ ДАННЫХ
НА РУССКОМ ЯЗЫКЕ**

**MAINTENANCE OF INFORMATION MANAGEMENT PROCESSES
USING A SOFTWARE
FOR AUTOMATED PROCESSING
OF SEMI-STRUCTURED TEXT DATA
IN RUSSIAN**

Макарова Е.А.,
Брянский Государственный Технический Университет,
Брянск, Российская Федерация

E.A. Makarova,
Bryansk State Technical University,
Bryansk, Russia

e-mail: m4karova.e@yandex.ru

Аннотация. Статья рассматривает использование слабоструктурированных текстовых данных в процессе информационного управления организационными системами. Описываются причины, по которым процесс информационного управления в экономических системах, где важной составляющей является постоянно обновляющаяся слабоструктурированная текстовая информация, целесообразно выстраивать по пути автоматизации, объединив сильные стороны машинной и ручной обработки. Представлена модель сбора и обработки слабоструктурированных текстовых данных при участии эксперта. Сформулированы функциональные требования к программному обеспечению для автоматизированной обработки слабоструктурированных текстовых данных на русском языке с привлечением эксперта, и разработан программный сервис, реализующий ряд методов и алгоритмов сбора и обработки текстовых данных. Описана архитектура веб-приложения, включающего в себя программный сервис, и структура подсистем программного сервиса и разработанной для него клиентской части. В качестве примера рассмотрены результаты применения предложенных решений к обработке новостного потока для анализа состояния в отрасли сельского хозяйства республики Татарстан. Данный анализ может быть полезен, например, для компаний, которые являются лизингодателями сельскохозяйственной техники. По результатам эксперимента уменьшено время на разведывательный анализ, что способствует более быстрому принятию решений, а в результате удаления дублей уменьшена выборка для предоставления эксперту.

Abstract. The paper considers the use of semi-structured textual data in the process of information management of organizational systems. It describes the reasons why the process of information management in economic systems, where an important component is constantly updated semi-structured textual information, is advisable to build along the path of automation, combining the strengths of machine and manual processing. A model for collecting and processing semi-structured text data with the participation of an expert is presented. Functional requirements for software for automated processing of semi-structured text data in Russian with the involvement of an expert have been developed, and a software service has been proposed that implements a number of methods and algorithms for collecting and processing text data. The architecture of a web application that includes a software service is described. The content of the subsystems of the software service and the client part developed for it has been disassembled. As an example, an analysis of such an area as monitoring the situation of agriculture in the Republic of Tatarstan is considered. This analysis can be useful, for example, for companies that are lessors of agricultural machinery. According to the results of the experiment, an decrease in the time for reconnaissance analysis was obtained, and as a result of the removal of duplicates, it is possible to reduce the sample to be provided to the expert.

Ключевые слова: обработка текстов, информационное управление, векторное представление слов, визуализация, программный сервис.

Keywords: text processing, information management, vector representation of words, visualization, software.

Использование в анализе слабоструктурированных источников данных, таких как, например, тексты, позволяет управленцам в экономических системах получать больше информации для принятия решений. При этом структура и формат необходимых данных может значительно различаться в зависимости от задачи. Исследуемые тексты могут отличаться по объему, структурированности, наличию профессиональной лексики и т.д. Многим текстам невозможно дать корректную оценку, для использования её в дальнейшем анализе, без привлечения эксперта в предметной области. В то же время, количество новых слабоструктурированных данных, появляющееся ежедневно, требует или привлечения большого количества ресурсов для их обработки [1], что весьма затратно, или автоматизации процесса.

Таким образом, процесс информационного управления в экономических системах, где важной составляющей является постоянно обновляющаяся слабоструктурированная текстовая информация, целесообразно выстраивать по пути автоматизации. Ранее совместно с соавторами были представлены различные подходы и методы автоматизации обработки текстовых данных для дальнейшего использования в принятии управленческих решений [2, 3, 4]. Все эти методы строятся на сочетании современных алгоритмов обработки естественного языка и привлечения эксперта с предоставлением ему интерактивных инструментов, сокращающих его время для решения задачи.

Для полноценной реализации предложенных методов необходимо программное обеспечение, реализующее все разработанные методы. Отдельные элементы предложенных методов реализованы в различных программных продуктах и библиотеках (PowerBI, WordStat, Gensim и т.д.), однако, необходима совместная реализация методов обработки текстов на серверной стороне и интерактивных интерфейсов для работы эксперта на клиентской стороне. Для этой цели разрабатывается программный сервис для автоматизированной обработки слабоструктурированных

текстовых данных на русском языке, предназначенный для поддержки процесса информационного управления.

Модель автоматизированной обработки слабоструктурированных текстовых данных

Общая модель автоматизированной обработки слабоструктурированных текстовых данных на русском языке в процессе информационного управления в динамических организационных системах представлена на рисунке 1.

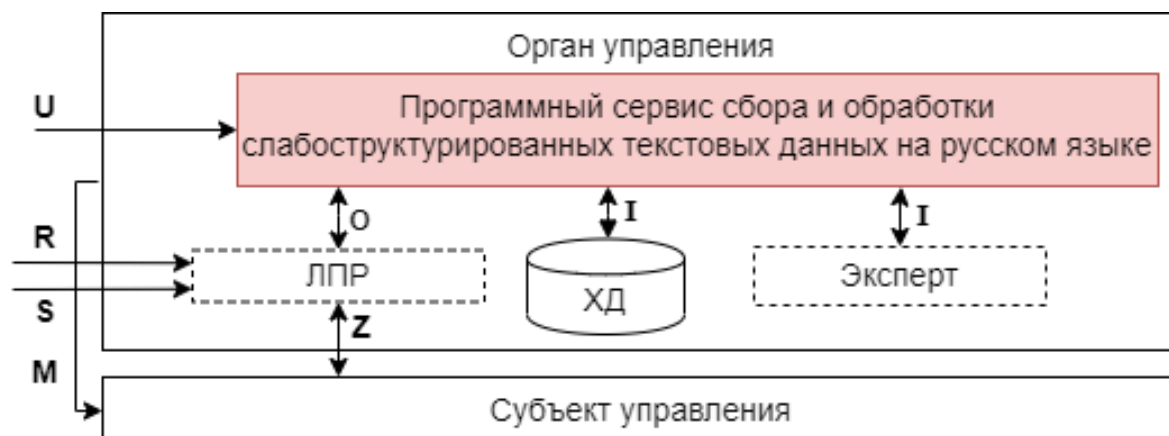


Рисунок 1. Модель автоматизированной обработки слабоструктурированных текстовых данных на русском языке в процессе информационного управления в организационных системах

Также данную модель можно описать следующим образом:

$$A = \langle R, U, M, D, Z; I \rangle,$$

где: R – задачи, которые необходимо выполнить лицу, принимающему решения;
 M – управленческие решения, непосредственно касающиеся объекта управления;
 D – информация о статусе объекта управления;
 Z – обратная связь по управленческому решению;
 I – информация, доступная для анализа и прогнозирования будущих состояний: выгруженные из источников данные и результат их автоматизированной обработки с привлечением эксперта;
 U – слабоструктурированные данные;
 S – структурированные данные;
 O – результаты обработки данных, информация для ЛПР.

Методы автоматизированной обработки информации будут зависеть от структуры информации и задач, которые необходимо выполнить лицу, принимающему решения. От корректности программной реализации разработанных методов будет зависеть оперативность получения информации, доступной для дальнейшего анализа (I). Для проектирования архитектуры программного сервиса необходимо сформулировать функциональные требования.

Основные функциональные возможности программного сервиса и клиентского приложения.

Основные функциональные возможности, предоставляемые сервисом, можно разделить на три этапа: сбор данных, их предобработка и выгрузка (экспорт) для дальнейшего использования в системах интеллектуального анализа данных или ручном анализе.

Функционал для сбора данных должен включать в себя:

- 1) настройка извлечения из источника, открытого или внутреннего, и подготовка данных для визуализации результатов извлечения
- 2) корректировка процесса извлечения с использованием результата работы эксперта с интерактивными визуализациями
- 3) выгрузка отдельных извлеченных документов

Функционал для обработки данных должен включать в себя:

- 1) настраиваемый процесс очистки данных
- 2) раскрытие специфических сокращений с привлечением эксперта в предметной области (предоставление ему интерфейсов для настройки и валидации процесса)
- 3) исправление ошибок с контролем эксперта
- 4) объединение текстов с высоким порогом семантической близости в группы для дальнейшего использования как одной сущности
- 5) разметка данных (ручная классификация) с предоставлением интерфейса с подсветкой семантически близких результатов

Функционал для выгрузки данных должен включать в себя выгрузку обработанных данных в одном из популярных форматов: JSON или xls.

Функциональность отдельных элементов, используемых в различных задачах, описана в предыдущих работах, посвященных отдельным методам.

Архитектура программного сервиса

На рисунке 2 представлена архитектура веб-приложения, включающего в себя программный сервис и клиентскую часть для работы с ним.

В подсистему сбора слабоструктурированных данных на русском языке входят:

- 1) модуль извлечения ключевых слов и группирования их по семантической близости;
- 2) модуль создания и взаимодействия с языковыми моделями (word2vec) [5];
- 3) модуль для формирования запросов для работы парсеров.

В подсистему автоматизированной обработки слабоструктурированных текстовых данных на русском языке входят:

- 1) модуль поиска и замены сокращений;
- 2) модуль исправления ошибочных написаний слов;
- 3) модуль определения семантической близости между текстовыми документами.

В клиентской части реализованы следующие компоненты:

- 1) интерактивная визуализация ключевых слов (словосочетаний) с вариантом реализации для людей с ограниченными возможностями зрения [6];
- 2) интерактивная визуализация моделей векторного представления слов;
- 3) редакторы для задач, связанных с обработкой и разметкой данных.

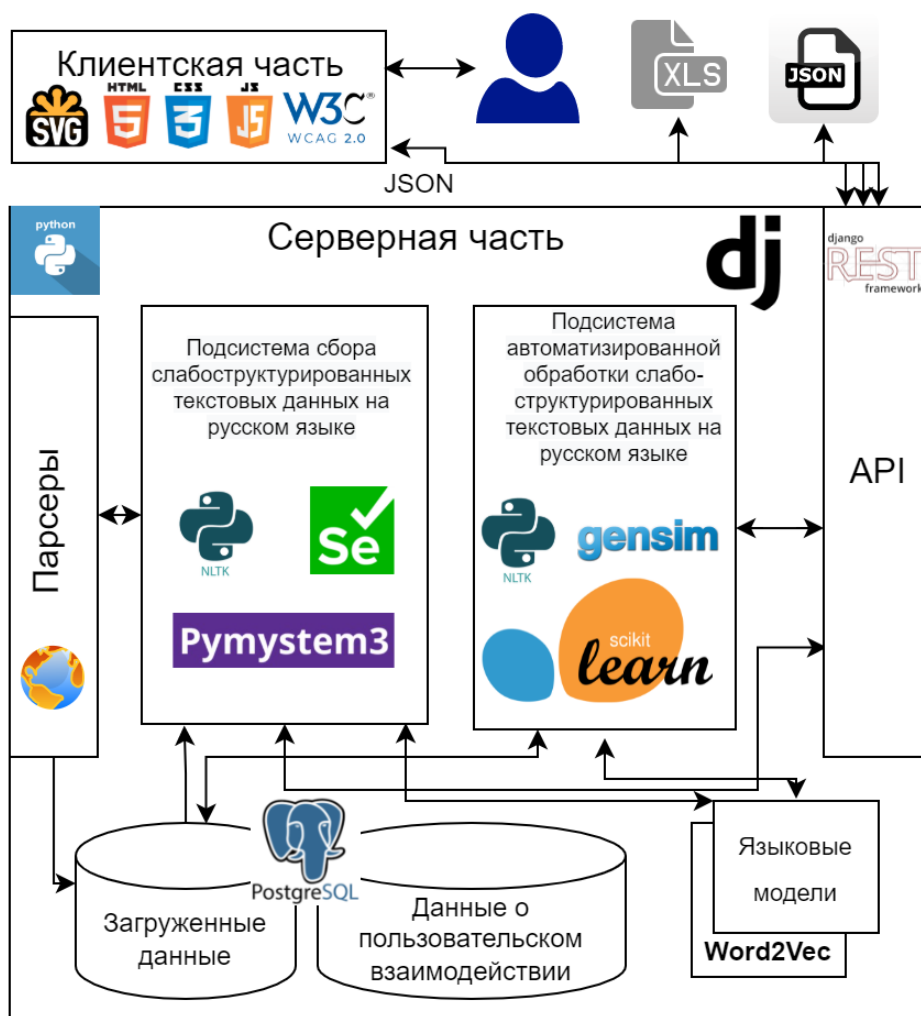


Рисунок 2. Архитектура веб-приложения

Результаты работы программного сервиса

Для проверки корректности работы программного сервиса он будет использоваться для задачи обработки новостного потока для информационного управления в динамических организационных системах. У решения подобной задачи могут быть следующие цели: анализ информации СМИ о текущем положении дел на предприятии, отрасли или регионе; мониторинг изменения трендов.

Программный сервис будет использоваться на всех этапах – от сбора до выгрузки обработанных данных. Данные для эксперимента взяты с крупнейших российских онлайн СМИ при использовании инструментов для сборки русскоязычных новостей и уже подготовленных сообществом баз проекта NewsViz (<https://github.com/newsviz>). В качестве примера рассмотрим анализ такой сферы как мониторинг положения сельского хозяйства в республике Татарстан. Данный анализ может быть полезен, например, для компаний, которые являются лизингодателями сельскохозяйственной техники. Для минимизации рисков, анализ положения дел в отрасли в конкретном регионе должен проводиться регулярно, чтобы не пропустить важных изменений тенденций.

Первый этап - настройка источников сбора данных и запроса. Так как нужную область характеризует много ключевых слов, пользователю необходимо будет ввести их все или пользоваться представленной на рисунке 3 интерактивной визуализацией. Для подбора синонимичных понятий используются модели word2vec – как «общие» для выбранного языка (включающего разные источники), так и «отраслевые» (в данном

случае, модель, обученная на корпусе новостей). На интерактивной визуализации пользователь может видеть степень семантической близости с нужным термином и выбрать необходимые. Таким образом, изначальное сочетание «Татарстан» + «сельскохозяйственный» превращается в более подробный набор слов для поиска информации. Информация о настроенном запросе сохраняется в базе данных.

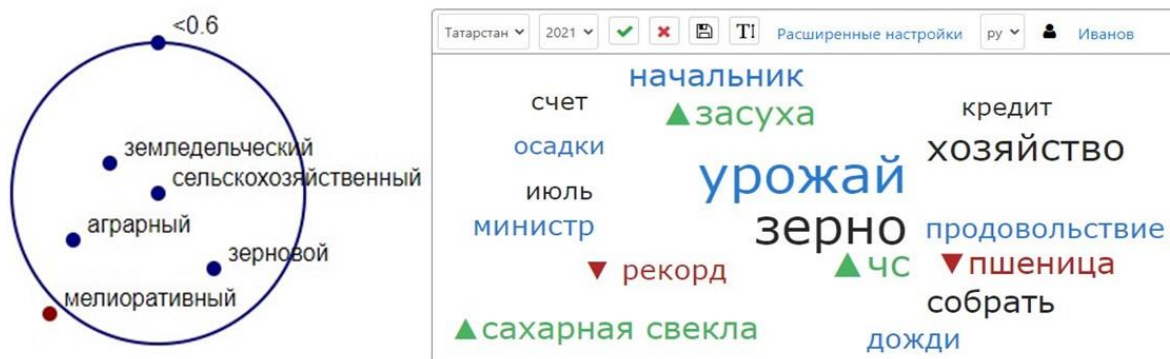


Рисунок 3. Интерактивная визуализация векторного представления слов и визуализация текстов с выделением трендов

Следующий этап в сборе данных – валидация процесса и настройка визуализации ключевых слов. При динамическом анализе (повторяющемся через определенные промежутки времени) данная визуализация отображает также появляющиеся тренды и значительное сокращение упоминаний. Результаты настройки сохраняются в базу данных, как и результаты предыдущего извлечения. Использование исторических данных позволяет выявить различные тренды. Так, например, если в 2020 году информационные ресурсы писали о рекордном урожае, то в 2021-ом ситуацию меняет засуха (рисунок 3).

Конечные выводы о влиянии описываемых в новостях событиях должен сделать эксперт в предметной области, сопоставив финансовые и прочие виды отчетности. Новостная повестка – лишь один из информационных входов, которые может учитывать ЛПР. Для достоверной оценки ситуации ему будет необходимо изучить не только динамику ключевых трендов, но и непосредственно тексты, описывающие возникшую ситуацию.

Для экономии времени на дальнейшую обработку и оценку данных из коллекции текстов необходимо удалить дубли. Подробнее алгоритм определения дублей описывался в работе [4]. Примеры степени семантической близости статей представлены в таблице 1.

Таблица 1 – Степень семантической близости заголовков

Заголовок 1	Заголовок 2	Степень близости
Татарстан в 2020г увеличил урожай зерна на четверть – до 5,3 млн тонн	Урожай зерна в Татарстане в 2020 году вырос до 5,3 млн. тонн	0.57
Урожай под угрозой: в Татарстане вводят режим ЧС из-за засухи	Более 40 районов Татарстана ввели режим ЧС из-за засухи	0.47

В результате удаления дублей возможно уменьшить выборку для предоставления эксперту до 20%. Выгрузка результатов работы может быть произведена в форматах xls или JSON. Выбор формата зависит от дальнейшего использования обработанной информации. Время на первоначальную настройку сервиса под задачи сравнимо с полностью ручным анализом (3-4 часа), однако, при повторном анализе, время значительно сокращается – до 0,5 часа, в зависимости от задачи.

Выводы

В процессе информационного управления экономическими системами целесообразно использовать слабоструктурированные текстовые данные. Но в ситуации частого обновления больших массивов данных, необходима автоматизация, которая объединит сильные стороны машинной и человеческой обработки.

Представленный в статье программный сервис позволяет использовать реализованные методы автоматизированной обработки текстов на русском языке для автоматизации процесса информационного управления.

Применение сервиса для анализа сообщений СМИ о положении дел в сельском хозяйстве республики Татарстан продемонстрировало возможность его применения для автоматизации обработки слабоструктурированных данных на русском языке. Благодаря интерактивным визуализациям был упрощен сбор данных, а алгоритмы для определения дублей уменьшили итоговую выборку документов для анализа ЛПР.

Некоторые модули сервиса ранее были использованы для проверки гипотез работоспособности разработанных методов для сбора и обработки слабоструктурированных текстовых данных на русском языке. Далее усилия планируется направить на нагрузочное тестирование разработанного сервиса и улучшение его функциональности по мере дальнейших исследований моделей обработки данных.

Литература

1. Pérez, J., Iturbide, E., Olivares, V. et al. A Data Preparation Methodology in Data Mining Applied to Mortality Population Databases//J Med Syst – 2015 – №39 – p. 152 – DOI: 10.1007/s10916-015-0312-5
2. Makarova E.A., Lagerev D.G., Lozbinev F.Y. Approaches to visualizing big text data at the stage of collection and preprocessing//Scientific Visualization – N. 4 – 2019.
3. Лагерев Д.Г., Макарова Е.А. Поиск и раскрытие сокращений в русскоязычных данных медицинских информационных систем//Вестник компьютерных и информационных технологий – № 7 – 2020 – с. 44-54.
4. Лагерев Д.Г., Макарова Е.А. Оценка семантической близости новостных сообщений на основе анализа заголовков//Вестник компьютерных и информационных технологий. – 2021. – Т. 18. – № 7(205). – С. 46-56. – DOI 10.14489/vkit.2021.07.pp.046-056.
5. Kutuzov A., Andreev I. Texts in, meaning out: neural language models in semantic similarity task for Russian//Proceedings of the Dialog. – Moscow – 2015.
6. Web Content Accessibility Guidelines (WCAG) 2.1. – Текст: электронный//The World Wide Web Consortium: [сайт]. – URL: <https://www.w3.org/TR/WCAG21/> (дата обращения: 29.03.2022).

УДК 004.9

**ПРОГРАММНЫЙ МОДУЛЬ РЕАЛИЗАЦИИ ШАБЛОНОВ
ДЛЯ СОЗДАНИЯ ДОКУМЕНТОВ
В СИСТЕМЕ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА**

**SOFTWARE MODULE FOR REALIZING TEMPLATES
FOR CREATING DOCUMENTS
IN THE ELECTRONIC DOCUMENT MANAGEMENT SYSTEM**

Муталлапов Р.Н., Шишкина О.Ю.,
Институт нефтепереработки и нефтехимии ФГБОУ ВО УГНТУ в г. Салавате,
ул. Губкина, 22 б, г. Салават, Республика Башкортостан,
Россия, 453250

Mutallapov R.N., Shishkina O.Yu.
Institute of Oil Refining and Petrochemistry FSBEI HE USPTU in Salavat,
Gubkin Str., 22b, Salavat, Republic of Bashkortostan,
453250, Russia

e-mail: olga79200@mail.ru

Аннотация. Качество управления организацией может оцениваться путем анализа уровня зрелости управления ее бизнес-процессов. Для оценки уровня зрелости управления бизнес-процессов используются разные модели. Наиболее используемой является модель «Business Process Management Maturity Model», которая включает в себя 5 уровней. Большинство российских компаний на данный момент находится на нулевом или первом уровне зрелости в соответствии с данной классификацией. Самым распространенным средством поддержки выполнения бизнес-процессов является документооборот, поэтому от качества ведения документооборота во многом зависит не только успех управленческой деятельности предприятия, но и его финансовое состояние. На основе проведенного анализа было выявлено, что внедрённая на предприятии система электронного документооборота реализована не до конца и имеет ряд недостатков. В данной статье рассматривается процесс «Документационное обеспечение» для внутренних документов предприятия. Для изучения предметной области был проведён детальный анализ бизнес-процессов «Документирование» и «Утверждение проекта документа», которые исходят из общего процесса. На основе проведённого анализа было выявлено, что в данных бизнес-процессах имеются проблемные области в процессе создания документа и утверждения документа электронной цифровой подписью (ЭЦП). Предполагается, что готовый модуль позволит сократить временные и материальные затраты, уменьшить количество ошибок, допущенных человеком при первичном заполнении документа.

Abstract. The quality of an organization's management can be assessed by analyzing the level of management maturity of its business processes. Different models are used to assess the maturity level of business process management. The most used is the Business Process Management Maturity Model, which includes 5 levels. Most Russian companies are currently at the zero or first level of maturity in accordance with this classification [1]. The most common means of supporting the implementation of business processes is document management, therefore, not only the success of the enterprise's management activities, but also its financial

condition largely depends on the quality of document management. Based on the analysis, it was revealed that the electronic document management system implemented at the enterprise has not been fully implemented and has a number of shortcomings. This article discusses the process of “Documentation” for the internal documents of the enterprise. To study the subject area, a detailed analysis of the business processes “Documentation” and “Approval of the draft document” was carried out, which come from the general process. Based on the analysis, it was revealed that these business processes have problem areas in the process of creating a document and approving a document with an electronic digital signature (EDS). It is assumed that the finished module will reduce time and material costs, reduce the number of mistakes made by a person when filling out a document for the first time.

Ключевые слова: система электронного документооборота, бизнес-процесс, шаблон, электронная цифровая подпись, программный модуль.

Keywords: electronic document management system, business process, template, electronic digital signature, software module.

Качество управления организацией во многом зависит от уровня зрелости управления бизнес-процессами организации. Под зрелостью управления понимается процесс управления организацией с помощью постоянного совершенствования методологии стратегического управления и ее интегрирования в общую систему управления. Одной из наиболее используемых моделей оценки уровня зрелости управления бизнес-процессов является «Business Process Management Maturity Model», которая включает в себя 5 уровней. Исходя из данной классификации большинство российских компаний на данный момент находится на нулевом или первом уровне зрелости [1].

Самым распространенным средством поддержки выполнения бизнес-процессов является документооборот, поэтому успех управленческой деятельности организации во многом зависит от качества ведения документооборота. Современные организации уже частично или полностью перешли на электронный документооборот, отказавшись от бумажного. Однако, если неправильно подойти к процессу перехода на электронный документооборот, можно нанести ущерб организации.

Сложность перехода на электронный документооборот во многом заключается в том, что в системах электронного документооборота (СЭД) необходима качественная настройка в соответствии с бизнес-процессами и регламентами компании. Проблемами работы в СЭД, которые не адаптированы под бизнес-процессы компании, являются: нарушение регламентов компании при создании документа, большой объем неэффективной ручной работы, создание документов с ошибками, при определенных условиях остаются затраты на бумажный документооборот.

Процесс «Документационное обеспечение» является одним из направлений автоматизации документооборота и внутренних бизнес-процессов компании.

Данный процесс состоит из пяти подпроцессов: «Документирование», «Обеспечение движения документации», «Утверждение проекта документа», «Контроль исполнения документации», «Организация хранения документации».

Рассмотрим процесс «Документирование» организационно-распорядительной и информационно-справочной документации (рисунок 1).

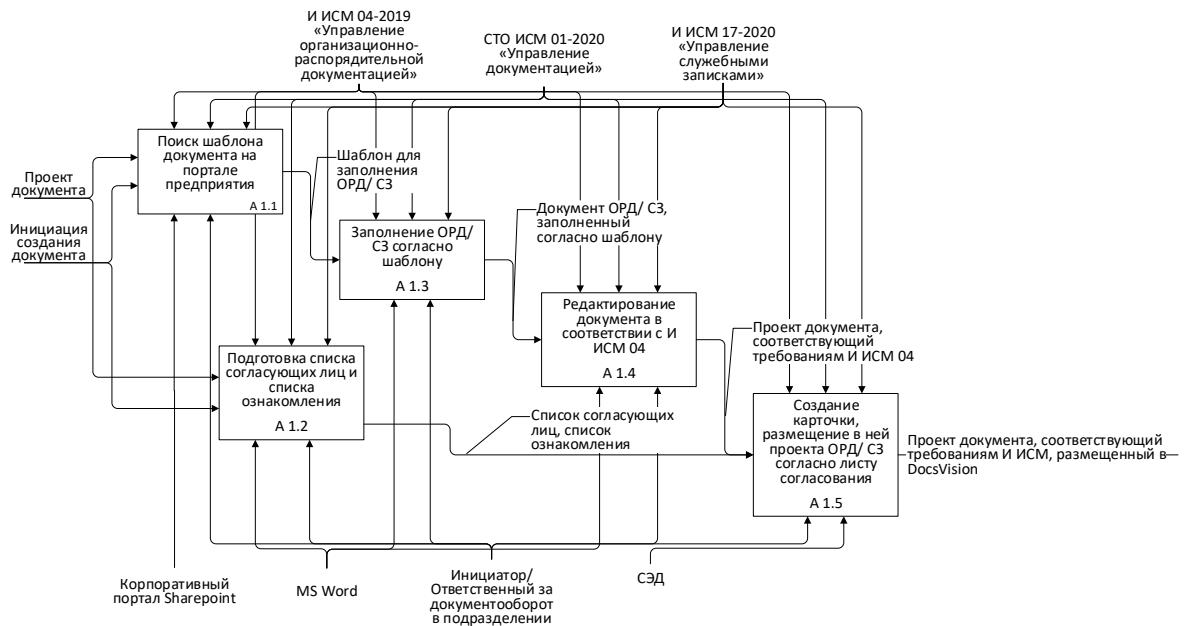


Рисунок 1. Диаграмма декомпозиции бизнес-процесса «Документирование»

Суммарная усредненная длительность бизнес-процесса «как есть» составляет 75 минут. В это время входит:

- Вход на портал предприятия, поиск шаблона документа, формы листа ознакомления, листа согласования (5 минут);
- Подготовка списка согласующих лиц и списка ознакомления (10 минут);
- Заполнение документа: ввод даты, констатирующей части, текста документа, должности и ФИО инициатора (30 минут);
- Редактирование документа в соответствии с И ИСМ 04 (10 минут);
- Создание карточки в СЭД, заполнение обязательных и необязательных полей документа, размещение проекта документа в СЭД согласно листу согласования (20 минут).

Рассмотрим процесс «Утверждение проекта документа» организационно-распорядительной и информационно-справочной документации (рисунок 2).

Суммарная усредненная длительность бизнес-процесса «как есть» составляет 79 минут. В это время входит:

- печать размещенного в СЭД проекта документа (2 минуты);
- утверждение документа живой подписью (40 минут);
- перевод документа в электронный вид (5 минут);
- загрузка в СЭД документа (2 минуты);
- рассылка согласно списку ознакомления (30 минут).

Каждый шаг бизнес-процесса рассмотрен на возможность возникновения риска,

- выделен объект риска,
- фактор риска,
- последствия, значимость риска (катастрофический, критический, существенный, граничный),
- вероятность наступления риска (очень высокая, довольно высокая, не слишком высокая, умеренная, небольшая, незначительная).

В таблице 1 представлен анализ рисков.

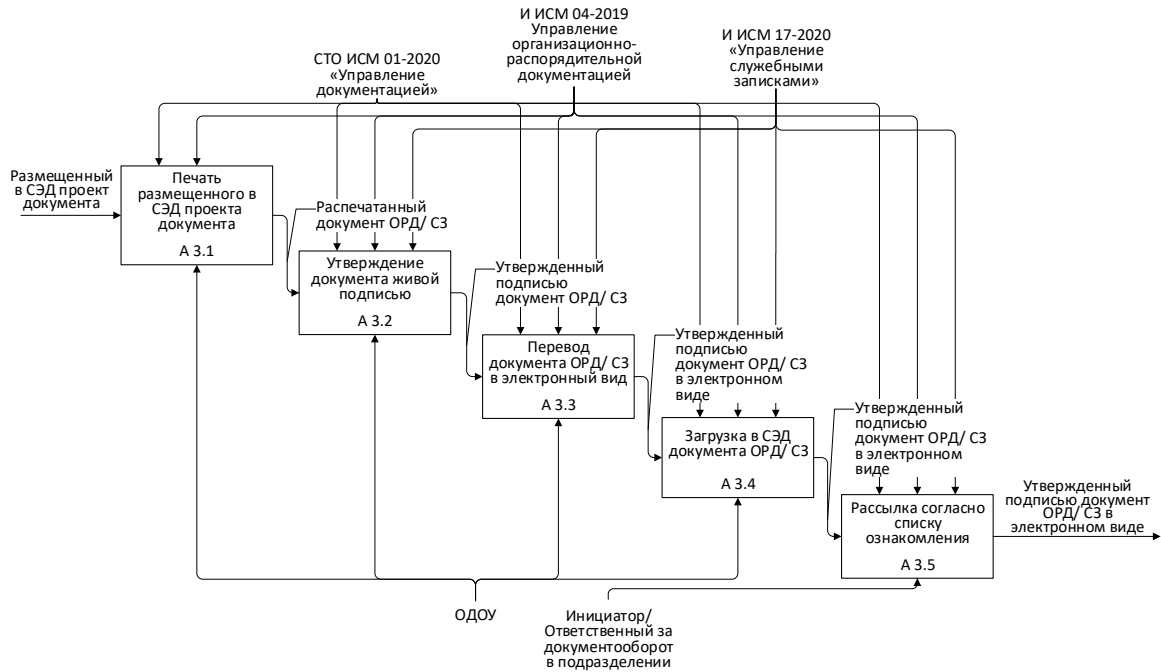


Рисунок 2. Диаграмма декомпозиции «Утверждение проекта документа»

Таблица 1. Анализ рисков

№	Объект риска	Фактор риска	Последствия	Значимость	Вероятность наступления риска
1	Поиск шаблона документа на портале предприятия	Ситуация, в которой невозможен доступ к порталу, нет единого шаблона	Невозможность выполнения проекта документа, большие затраты времени	Существенная	Умеренная
2	Заполнение документа проекта ОРД/СЗ	Ошибка при заполнении, нарушение/неисполнение регламентов	Большие затраты времени	Существенная	Не слишком высокая
3	Подготовка списка согласующих лиц и списка ознакомления	Ошибка при заполнении	Убытки в результате принятия ошибочных решений	Критическая	Незначительная
4	Редактирование документа в соответствии со стандартами	Ошибка при заполнении	Большие затраты времени	Граничный	Небольшая
5	Согласование проекта документа, направление по маршруту	Ошибка при проверке документа, заполнении карточки в СЭД	Возникновение убытков в результате принятия ошибочных решений	Существенный	Незначительная

Продолжение таблицы 2

6	Печать размещенного в СЭД проекта документа	Потеря, утрата документа	Возникновение убытков; затрачивание большего количества времени	Граничный	Довольно высокая
7	Утверждение документа живой подписью	Потеря, утрата документа	Возникновение убытков; большие затраты времени	Граничный	Довольно высокая
8	Перевод документа ОРД/ СЗ в электронный вид	Качество работы периферии	Возникновение убытков; большие затраты времени	Существенный	Умеренная
9	Загрузка в СЭД ОРД/ СЗ	Ошибка в составлении карточки в СЭД	Возникновение убытков; большие затраты времени	Существенный	Небольшая
10	Рассылка согласно списку ознакомления	Ошибка при заполнении	Большие затраты времени	Граничный	Незначительная

После идентификации и оценки рисков необходимо разработать меры, обеспечивающие минимизацию вероятности и ослабление отрицательных последствий рисков событий.

Создание программного модуля реализации шаблонов для создания документов в системе электронного документооборота позволит снизить такие риски как:

- ошибки при заполнении документа проекта ОРД/ СЗ, с помощью проверки введенного текста на орфографию и пунктуацию. Такую меру можно отнести к стратегии снижения – вероятность реализации риска снижена, последствия негативного рисковог события снижены до приемлемых пределов – риск либо не сбудется, либо сбудется, но с меньшими последствиями.

- ошибки при заполнении на этапе редактирования документа в соответствии со стандартами, с помощью формирования текста по готовому шаблону, что исключает опечатки. Такую меру также можно отнести к стратегии снижения риска.

- потеря, утрата документа при печати размещенного в СЭД проекта документа, утверждении документа живой подписью, с помощью возможности подписания документа электронной цифровой подписью. Такая мера относится к стратегии уклонения – исключает угрозу.

- качество работы периферии при переводе документа ОРД/ СЗ в электронный вид, так же с помощью возможности подписания документа электронной цифровой подписью. Применяется стратегия уклонения.

- ошибка в составлении карточки в СЭД (прикрепления другого документа) при загрузке в СЭД ОРД/ СЗ с помощью генерации созданных документов в формат docs, pdf и добавлении сгенерированного файла в карточку. Применяется стратегия снижения.

С учетом карты рисков построена новая модель бизнес-процессов «Документирование» и «Утверждение проекта документа».

Бизнес-процесс «Документирование» исключает распараллеливание выполняемых задач, становится последовательным. На рисунке 3 также показана функция, которую реализует программный модуль реализации шаблонов для создания документов в системе электронного документооборота – генерация в MS Word/формирование СЗ. Таким образом рассматриваемый модуль выступает инструментом для создания документа в системе электронного документооборота.

Программный модуль позволит исключить такие подпроцессы как: поиск шаблона документа на портале предприятия, заполнение ОРД/ СЗ согласно шаблону, размещение проекта документа ОРД/ СЗ в карточке СЭД, изменить такие процессы как: подготовка списка согласующих лиц и списка ознакомления – эта информация вводится в карточку документа при ее создании, редактирование документа в соответствии с И ИСМ 04 – частичное редактирование будет производиться автоматически, за исключением таблиц.

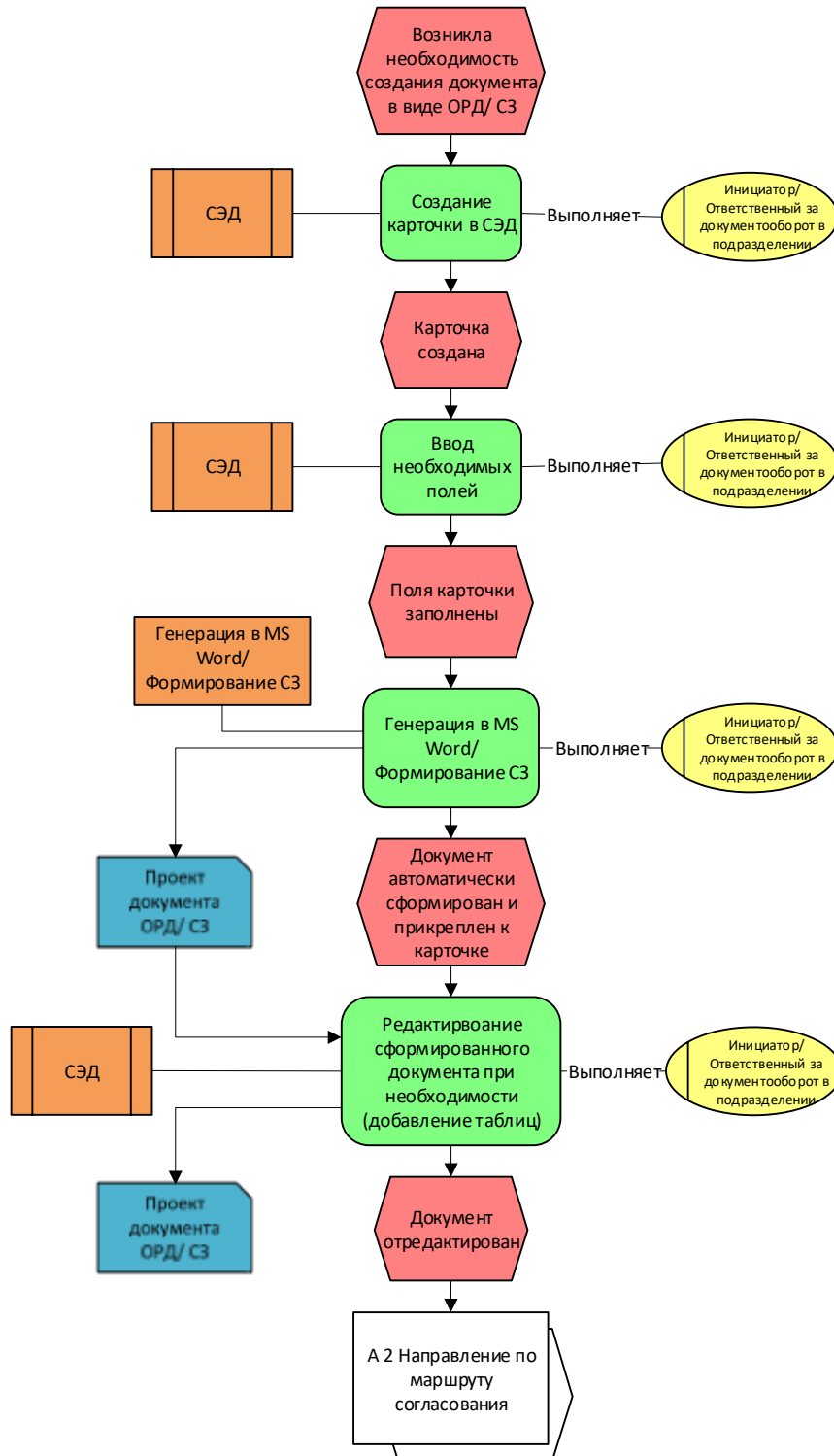


Рисунок 3. Диаграмма декомпозиции «Документирование» «как будет»

На рисунке 4 представлен бизнес-процесс «Утверждение проекта документа», в котором также показана функция, которую реализует программный модуль реализации шаблонов для создания документов в системе электронного документооборота – подписание проекта документа ЭЦП. Таким образом рассматриваемый модуль выступает инструментом для подписанта по утверждению проекта документа.

Программный модуль позволит исключить такие подпроцессы как: печать проекта документа, перевод документа ОРД/ СЗ в электронный вид, загрузка в СЭД документа ОРД/ СЗ, изменить такой процесс как утверждение документа живой подписью – с помощью появления возможности подписания документа ЭЦП. За счет этого программный модуль позволит исключить материальные затраты на бумажный документооборот (бумага, картридж, ремонт многофункциональных устройств и пр.).

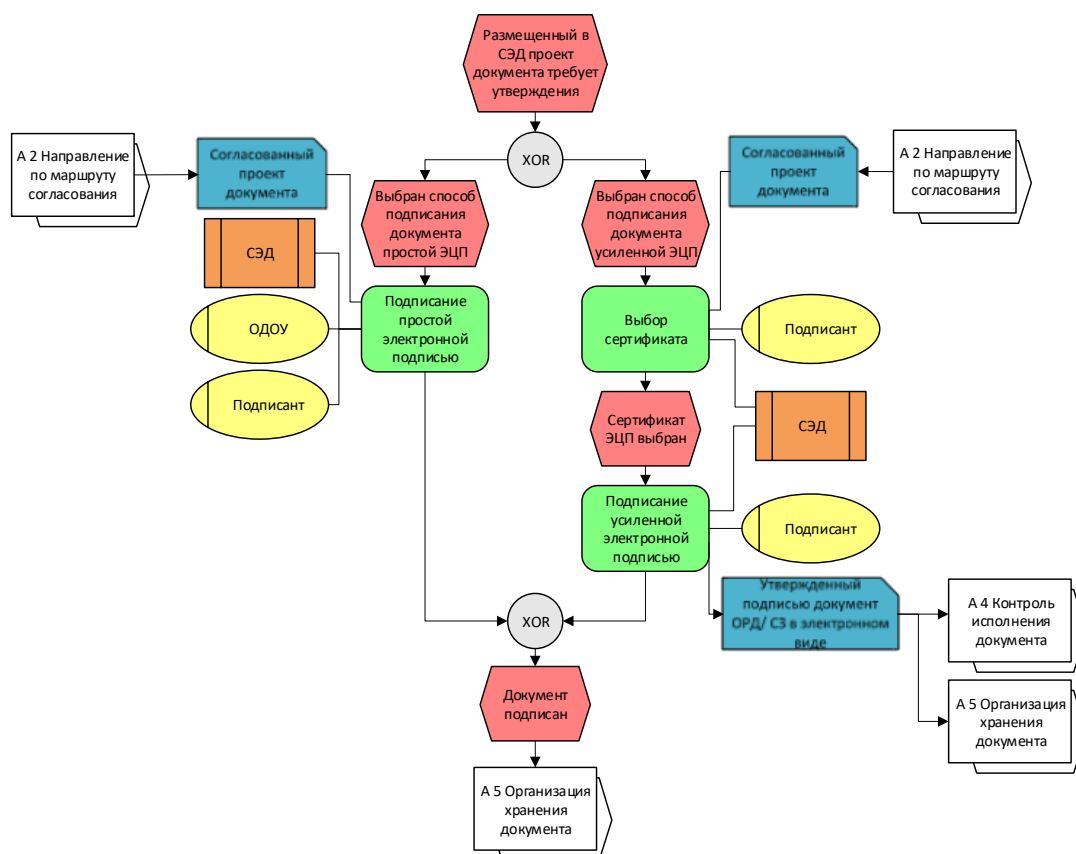


Рисунок 4. Диаграмма декомпозиции «Утверждение проекта документа» «как будет»

Выводы

Использование программного модуля реализации шаблонов для создания документов в системе электронного документооборота позволит сократить временные и материальные затраты, уменьшить количество ошибок, допущенных человеком при первичном заполнении документа.

Литература

1. Долганова, О.И. Моделирование бизнес-процессов [Текст]: учебник и практикум для академического бакалавриата/О.И. Долганова, Е.В. Виноградова,

А.М. Лобанова; под ред. О.И. Долгановой. – М.: Изд-во Юрайт, 201. – 289 с. – Серия: Бакалавр. Академический курс.

2. Головина, Е.Ю. Сравнение архитектур систем электронного документооборота/Е.Ю. Головина, Е.В. Ермолаев//Информационные технологии. Проблемы и решения. – 2020. – № 3(12). – С. 41-46.

3. Левина Т.М. Применение информационных технологий при рассмотрении экономических показателей в управлении проектами/Т.М. Левина, Н.Н. Лунева, К.В. Галиуллина//Информационные технологии. Проблемы и решения: материалы Междунар. науч.-практ. конф.: дополнит. сб. -Уфа: Изд-во УГНТУ, 2014. – С. 239-243.

УДК 004

ПРИМЕНЕНИЕ СИСТЕМЫ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА В ЮРИДИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ ОРГАНИЗАЦИИ

APPLICATION OF THE ELECTRONIC DOCUMENT MANAGEMENT SYSTEM IN THE LEGAL ACTIVITY OF THE ORGANIZATION

¹Мовсесян У.А., ¹Ткаченко А.Л., ²Федорова В.А.,

¹Калужский государственный университет им. К.Э. Циолковского,
г. Калуга, Российская Федерация

²Московский государственный технический университет им. Н.Э. Баумана,
г. Москва, Российская Федерация

U.A. Movsesyan¹, A.L. Tkachenko¹, V.A. Fedorova²,
¹Kaluga State University named after K.E. Tsiolkovsky,
Kaluga, Russian Federation

²Bauman Moscow State Technical University,
Moscow, Russian Federation

e-mail: movsesyanua@studklg.ru

Аннотация. Во втором десятилетии 21 века более очевидным становится решающая роль информатизации в управлении персоналом, происходит быстрый перехода к компьютерным технологиям. В юридической сфере использование компьютерных технологий позволяет увеличить скорость, качество и достоверность протекающих процессов. Актуальность данной темы обусловлено тем, что процесс информатизации способствует развитию юридической науки и в теоретическом и практическом аспекте, способствует изучению предпосылок, проблемных вопросов и перспектив развития юридической деятельности в современных условиях. Данный вопрос по настоящее время не потеряла свою актуальность, так как документооборот является основой в организации работы с документами. Документооборот предприятия - это комплекс системных процедур, гарантирующий движение отчетов в компании от начала и до окончания исполнения конечным пользователем. Для правильной и целесообразной оптимизации документооборота все отчеты распределяются по документопотокам. Качественной особенностью документооборота представляется количество и объем информации, проходящий через него. Это число отчетов, поступивших в предприятие и сформированных ею в установленные сроки, как показывает практика, в течении одного года. Объем документооборота – один из главных

характеристик, играющая решающую роль при принятии выбора электронной системы документооборота.

Abstract. In the second decade of the 21st century, the decisive role of informatization in personnel management becomes more obvious, there is a rapid transition to computer technology. In the legal field, the use of computer technology allows you to increase the speed, quality and reliability of ongoing processes. The relevance of this topic is due to the fact that the process of informatization contributes to the development of legal science in both theoretical and practical aspects, contributes to the study of prerequisites, problematic issues and prospects for the development of legal activity in modern conditions. The document flow of an enterprise is a set of system procedures that guarantees the movement of reports in the company from the beginning to the end of execution by the finished user. For proper and expedient optimization of document flow, all reports are distributed across document flows. A qualitative feature of the document flow is the amount and volume of information passing through it. This is the number of reports received by the company and generated by it on time, as practice shows, within one year. The volume of document flow is one of the main characteristics that plays a decisive role in making the choice of an electronic document management system.

Ключевые слова: система электронного документооборота, информационные технологии, юридическая деятельность, «Тезис», передача информации.

Keywords: electronic document management system, information technologies, legal activity, “Thesis”, transmission of information.

В юридической сфере часто возникает необходимость использовать разного рода нормативно-правовые акты. В ходе своей работы всякий юрист основывается на Конституции Российской Федерации, международных договорах Российской Федерации, федеральных конституционных законах, федеральных законах, нормативных правовых актах Президента РФ, нормативных правовых актах Правительства РФ, нормативных правовых актах федеральных органов государственной власти, конституциях, законах, иных нормативных правовых актах органов государственной власти субъектов Российской Федерации, нормативных правовых актах органов местного самоуправления [1-3].

Полная и достоверная информация лежит в основе любого решения юриста. Немаловажную роль играет удобство поиска, хранения и применения информации. Система легкого поиска, оформления и передачи документов оптимизирует работу юристов, поскольку юрист имеет возможность экономить время на технические вопросы и рутинную работу и сконцентрироваться на содержательной и созидательной деятельности. Электронная система документооборота может сыграть важную роль в такой сложной сфере как юриспруденция, с огромным количеством информации и документов [4-7].

Электронный документооборот дает возможность выполнять все те же задачи, что и его аналог на бумаге. Он позволяет создавать, изменять, регистрировать документы, отправлять их адресатам. Так же, электронный документооборот сильно облегчает формирование и отправку отчетов в Федеральную Налоговую Службу или другие органы контроля. Электронные документы должны быть годны для проверки разного рода системами. Для того, чтобы отправленный отчет имел достаточное юридическое значение, он подписывается электронной подписью.

В данной ситуации более актуальным становится использование новых систем электронного документооборота, одним из таких систем является СЭД «ТЕЗИС».

СЭД «ТЕЗИС» – это система для организации электронного документооборота внутри компании. Пользователи программы могут настроить обмен документами между отделами и филиалами, контролировать весь путь документа до конечного адресата – регистрацией, обработкой, исполнением, хранением.

Работать с документами в СЭД помогут шаблоны для автозаполнения, откат редакций файла, возврат на доработку, передача в архив и извлечение из него. Документы автоматически нумеруются и могут быть разделены на тома.

Главные функциональные системы – управление задачами, документами и канцелярией (рисунок 1).

Подтверждение
✕

Детали
Вложения

+ Добавить ▾
🗑

	Роль	Пользователь	E-mail	Очередность	
	Инициатор	Кошелев Н. А. ▾	☑		🗑
+	Исполнитель	Сидоров Б. Ю. 👤 ▾	☑	1 ▾	🗑
+	Согласующий	Петров Б. Ю. 👤 ▾	☑	1 ▾	🗑
+	Согласующий	Тяпкина Р. О. 👤 ▾	☑	1 ▾	🗑
	Утверждающий	Иванов К. М. 👤 ▾	☑		🗑

Комментарий

Введите текст

Завершить к

ДД.ММ.ГГГГ
📅

ЧЧ.ММ.
🕒

Отправить в ЭДО

Отменить
Подтвердить

Рисунок 1. Подтверждение списка пользователей

Управление задачами:

- назначение задач исполнителю, управление исполнением задач;
- постановка задач одному или нескольким исполнителям, контроль исполнения задач;
- создание различных видов задач: зависящих друг от друга, подчиненных, периодических, отложенных, относящихся к протоколу совещания, а также создание групп задач;
- импорт задач из Microsoft Project;
- формирование отчетов в формате .xls;
- графическое представление поставленных/исполняемых задач в виде диаграммы Ганта;
- организация служб техподдержки HelpDesk/ServiceDesk;

Управление документами:

- автоматизация связанных с документами процессов: создание, согласование, ознакомление, утверждение;
- наглядное представление соподчиненности документов, контроль обновлений документов;
- просмотр документов с помощью Microsoft Office Web Apps;
- наглядное отображение маршрута движения документов в виде диаграммы.

Электронная канцелярия:

- учет входящих, исходящих и внутренних документов;
- механизм предотвращения повторной регистрации входящего документа;
- наложение резолюции на документ и создание задачи на его основании;
- создание номенклатуры дел и настройка нумерации дел и документов в соответствии с политикой организации-заказчика;
- печать конвертов для документов и договоров;
- отслеживание истории переписки с контрагентом;
- организация электронного архива документов;
- потоковое сканирование и обработка документов.

Для полного представления и контроля стадии дел в организации, система электронного документооборота «ТЕЗИС» предоставляет возможность получать актуальную информацию о ходе выполнении задач сотрудниками, что позволяет улучшать оперативность и эффективность принимаемых управленческих решений, а также повышает работоспособность и качество работы персонала, что способствует снижению рисков потерь.

Каждый сотрудник организации, получающий задачи и поручения через систему электронного документооборота, имеет доступ к шаблону автоматического заполнения отчета, на основе ранее выполненных работ. Это значительно сокращает время на создание документа, а также дает возможность информировать руководство о реальной загрузке и трудоёмкости хода выполнения работы.

Работа в системе документооборота «ТЕЗИС» позволяет руководителю экономить время на рутинных обязанностях по отслеживанию исполнения своих поручений при помощи автоматизации данной системы, что позволяет уделять больше времени и внимания на контроль качества исполненных задач (рисунок 2).

Кроме того, в дополнение к основным возможностям системы электронного документооборота, «ТЕЗИС» облегчает выполнение разного рода задач делопроизводителя и открывает новые перспективы для самостоятельного расширения системы внутри организации.

В настоящее время, на отечественном рынке систем электронного документооборота, СЭД «ТЕЗИС» конкурирует с такими программами как 1С-ЭДО и ЕВФРАТ.

Они находятся в одной ценовой категории и имеют схожие функциональные возможности.

На пользовательском рынке все три программы имеют ту же целевую аудиторию. Данные системы подходят для разного вида государственных учреждений, предприятий и коммерческих компаний, чтобы автоматизировать электронный документооборот.

Подробное сравнение функционала данных систем электронного документооборота рассмотрено в таблице 1.

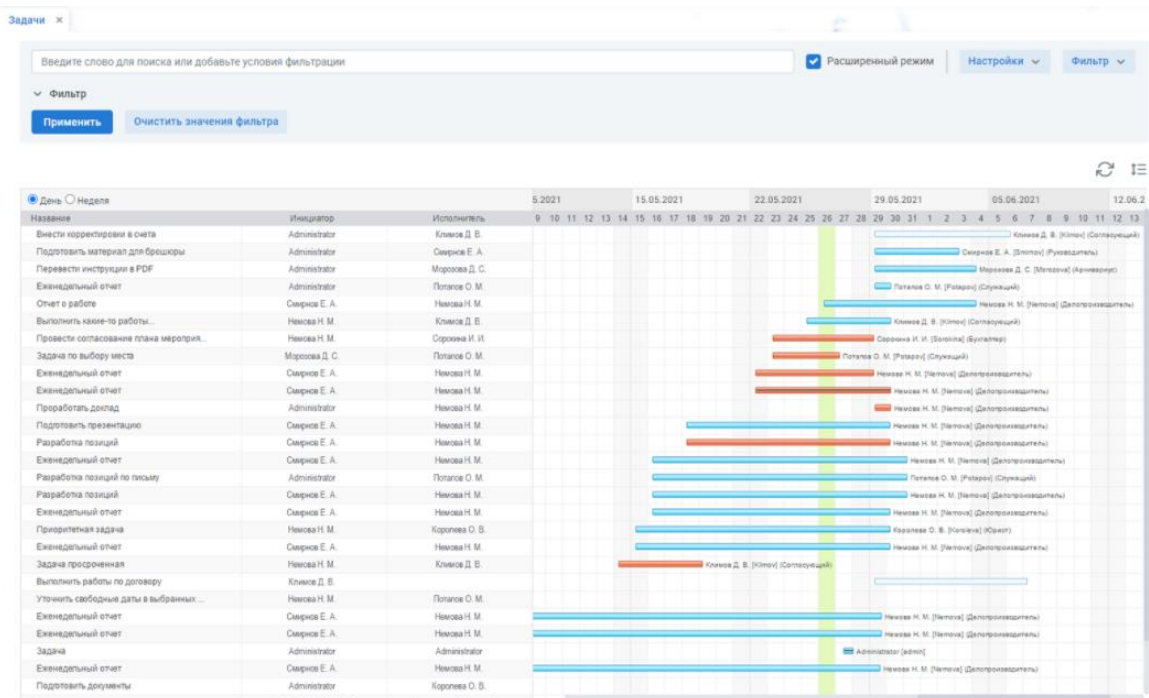


Рисунок 2. График контроля исполнения

Таблица 1 – Сравнение СЭД ТЕЗИС с 1С-ЭДО и ЕВФРАТ

Функционал СЭД	ТЕЗИС	1С-ЭДО	ЕВФРАТ
Стартовая стоимость базовых тарифов	От 35 тыс. руб.	От 10 тыс. руб.	От 27 тыс. руб.
Автоматизация рабочих процессов	Да	Нет	Да
Работа с текстом документа в виде вложения	Да	Нет	Да
Наличие возможности фильтрации и сортировки списков документов	Да	Да	Да
Хранение документооборота (электронный архив)	Да	Нет	Да
Механизм гарантированной доставки	Нет	Да	Нет
Контроль доступа/разрешения	Да	Нет	Нет
Наличие внутри системы календаря для просмотра задач	Да	Да	Нет
Экспорт документов в форматы (pdf, doc)	Да	Нет	Да
Полнотекстовый поиск	Да	Нет	Нет
Сотрудничество	Да	Нет	Нет
Открытый код	Нет	Да	Нет
Управление версиями	Да	Нет	Нет
Мобильные платформы IOS и Android	Да	Нет	Нет
Оповещения на почту	Да	Да	Да
Инструкция пользователю	Да	Да	Да

Выводы

Анализ функциональных возможностей и характеристик вышеуказанных систем электронного документооборота показывает, что СЭД «ТЕЗИС» предлагает больше возможностей и является лучшим решением, несмотря на более высокую стоимость.

Литература

1. Управленческий документооборот: от бумажного к электронному. Вопросы теории и практики/М.П. Бобылева//ТЕРМИКА. – 2019. – С. 110-116.
2. Ткаченко, А. Л. Анализ и рекомендации по выбору аналитической платформы /А.Л. Ткаченко, И.А. Лыгин, В.И. Кузнецова//Заметки ученого. – 2021. – № 7-1. – С. 51-54.
3. Испирян, Р.А. Принятие управленческих решений посредством системы технического диагностирования/Р.А. Испирян, А.Л. Ткаченко//Развитие управленческих и информационных технологий, их роль в региональной экономике: материалы II Международной открытой научно-практической конференции, Калуга, 21-22 апреля 2016 года/Под редакцией: Пироговой Т.Э., Швецовоy С.Т., Орловцевоy О.М. – Калуга: ООО «ТРП», 2016. – С. 95-102.
4. Павлюк, А.Я. Системы электронного документооборота и управление отношениями с клиентами/А.Я. Павлюк, А.Л. Ткаченко//Актуальные вопросы современной науки: сборник статей по материалам XVIII международной научно-практической конференции, Томск, 13 февраля 2019 года. – Томск: Общество с ограниченной ответственностью Дендра, 2019. – С. 95-99.
5. Кондрашова, Н.Г. 4.1. Риск-ориентированный внутренний контроль: практическая реализация/Н.Г. Кондрашова//Аудит и финансовый анализ. – 2019. – № 2. – С. 60-64.
6. Ткаченко, А.Л. Имитационное моделирование распространения кибератак на промышленные предприятия/А.Л. Ткаченко, А.Ю. Гордеева, А.В. Шавренко//Инновационные технологии, экономика и менеджмент в промышленности: Сборник научных статей по итогам IV международной научной конференции, Волгоград, 22-23 апреля 2021 года. – Волгоград: Общество с ограниченной ответственностью «КОНВЕРТ», 2021. – С. 238-240.
7. Ткаченко, А.Л. Использование программы anylodgic для анализа потребительского рынка игры Genshin impact и построение прогноза/А.Л. Ткаченко, О.И. Ольшанская, О.Н. Арышева//Информационные технологии. Проблемы и решения. – 2021. – № 4(17). – С. 46-51.

МОДЕЛИРОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ

УДК 004.6

**МОДЕЛИ И АЛГОРИТМЫ СБОРА ИНФОРМАЦИИ
ДЛЯ ФОРМИРОВАНИЯ БАЗЫ ДАННЫХ
СИСТЕМЫ МОНИТОРИНГА ЛЕСНЫХ ПОЖАРОВ****MODELS AND ALGORITHMS
FOR COLLECTING INFORMATION TO FORM THE DATABASE
OF THE FOREST FIRE MONITORING SYSTEM**

Лапин А.Н., Миначов Ш.М., Широкова А.А.,
ФГБОУ ВО «Уфимский государственный авиационный технический университет»,
г. Уфа, Российская Федерация

A.N. Lapin, S.M. Minasov, A.A. Shirokova,
Ufa State Aviation Technical University,
Ufa, Russian Federation

e-mail: meccos160@yandex.ru

Аннотация. В статье описана технология сбора и хранения официально опубликованных данных о лесных пожарах, произошедших на территории Российской Федерации за последние годы. В работе проведен анализ современного состояния предметной области, проведена оценка актуальности поставленной проблемы. По результатам анализа выявлены проблемы, возникающие при тушении лесных пожаров на удалённых от мест дислокации сил МЧС территориях, сформулирована гипотеза о целесообразности ведения мониторинга территории пожара и создания его цифрового двойника. Собрана информация из открытых источников данных о лесных насаждениях. Произведён анализ данных о лесах в регионах РФ и представлена традиционная для страны геоботаническая типология. Приведена классификация лесных пожаров по скорости распространения и характеру возгорания. Изучена структура публикации данных о лесных пожарах в открытых источниках, в частности на сайте Федерального бюджетного учреждения «Центральная база авиационной охраны лесов». На основе выполненного анализа разработан типовой алгоритм и реализована программа для извлечения данных со страниц указанного сайта. Разработана информационная модель системы, на основе которой создана база данных для хранения извлекаемой из открытых источников информации. Разработанные модели и алгоритмы позволили эффективно собрать данные, для последующей реализации в системе мониторинга и прогнозирования распространения фронта лесных пожаров.

Abstract. The article describes the technology for collecting and storing officially published data on forest fires that occurred on the territory of the Russian Federation in recent years. The paper analyzes the current state of the subject area and assesses the relevance of the problem posed. By results of the analysis the problems arising at suppression of forest fires on territories remote from places of deployment of forces of the Ministry of Emergency Measures are revealed, the hypothesis about expediency of conducting monitoring of territory of fire and creation of its digital duplicate is formulated. Information from open sources of data on forest

plantations was collected. The analysis of data on forests in regions of the Russian Federation is made and the geobotanical typology traditional for the country is presented. The classification of forest fires by speed of spread and nature of ignition is given. The structure of publication of data on forest fires in open sources, in particular on the site of the Federal Budget Institution “Central base of aviation forest protection” was studied. On the basis of the analysis a typical algorithm was developed and a program to extract data from the pages of the site was implemented. The information model of the system has been developed and used to create a database to store information retrieved from open sources. The developed models and algorithms allowed us to effectively collect data for further implementation in the system of monitoring and forecasting the spread of forest fires.

Ключевые слова: лесной пожар, мониторинг, парирование нештатной ситуации, база данных, реляционная модель, СУБД, обработка данных в реальном масштабе времени.

Keywords: forest fire, monitoring, emergency response, database, relational model, DBMS, real-time data processing.

Введение

Лес является одним из важнейших природных богатств, которое обеспечивает жизнедеятельность природы и человека. На территории Российской Федерации располагается четверть мировых запасов леса [1]. По данным Единой межведомственной информационно-статистической системы, площадь лесных земель России, пройденная пожарами, лишь за последний квартал 2021 года, составляет более 8 млн. га. А всего за неделю, с 7 по 13 марта 2022 года, по информации региональных диспетчерских служб лесного хозяйства, общая площадь, пройденная огнём, составляет более 3 тысяч га [2].

Сокращение площади территорий распространения пожаров и минимизация ущерба от стихийных возгораний обеспечивается соответствующими мерами, например, изготовлением в лесных массивах, вокруг застроенных жилых и производственных территорий противопожарные минерализованные полосы, высадки не распространяющих огонь сельскохозяйственных культур и прочих мер. Но такие меры невозможно либо слишком дорого осуществлять по принципу «давайте сделаем везде на всякий случай».

С учетом вышесказанного для принятия верного решения в отношении необходимости проведения противопожарных мероприятий, а также разработки тактики тушения конкретных возгораний, необходимо иметь возможность прогнозирования ситуаций, которая должна быть основана на соответствующих данных и результатах их анализа.

В настоящее время наряду с официальными данными, которые не всегда содержат полноценную информацию, необходимую для корректного моделирования, существует множество открытых источников информации о пожарах на территории РФ, но данные в этих источниках расположены не структурированно, что не позволяет их эффективно использовать.

Анализ рынка прикладного ПО также показал отсутствие реализованных решений по мониторингу лесных пожаров в удаленных районах в режиме реального времени. Таким образом, создание системы мониторинга и прогнозирования распространения лесных пожаров является актуальной проблемой.

Целью настоящего проекта является повышение эффективности процесса принятия решений для минимизации ущерба стихийных лесных пожаров и затрат на их тушение на удалённых от служб МЧС территориях.

Для достижения указанной цели поставлены следующие задачи:

- выполнить сбор, структурирование и анализ данных о лесных насаждениях РФ;
- разработать структуру базы данных для накопления и обработки данных о лесных пожарах;
- разработать алгоритмы автоматического сбора из открытых источников;
- выполнить сбор и анализ информации о стихийных лесных пожарах, произошедших на территории РФ за последние 20 лет;
- проанализировать существующие методы прогнозирования распространения лесных пожаров;
- выполнить оценку эффективности предложенных решений.

Классификация лесов и их распространение в регионах РФ

В Российской Федерации леса классифицируются [3]:

- по преобладающим породам (еловые, сосновые, лиственничные, дубовые, берёзовые и др.)
- в зависимости от богатства и увлажнения почв (например, сосняки лишайниковые, брусничные, кисличные, сфагновые и т.д.).

В.Н. Сукачёв [4] предложил биогеоценотический подход к классификации лесов, согласно которому характерная черта леса – его динамичность, которая определяется, с одной стороны, влиянием внешних условий, а с другой – автогенными причинами, обусловленными жизнедеятельностью самого леса (рисунок 1).



Рисунок 1. Геоботаническая типология леса В.Н. Сукачёва

На рисунке 2 приведена диаграмма распространения хвойных лесов на территории РФ. В [5] также приведены подробные данные об общих запасах твердолиственных и мягколиственных лесов.

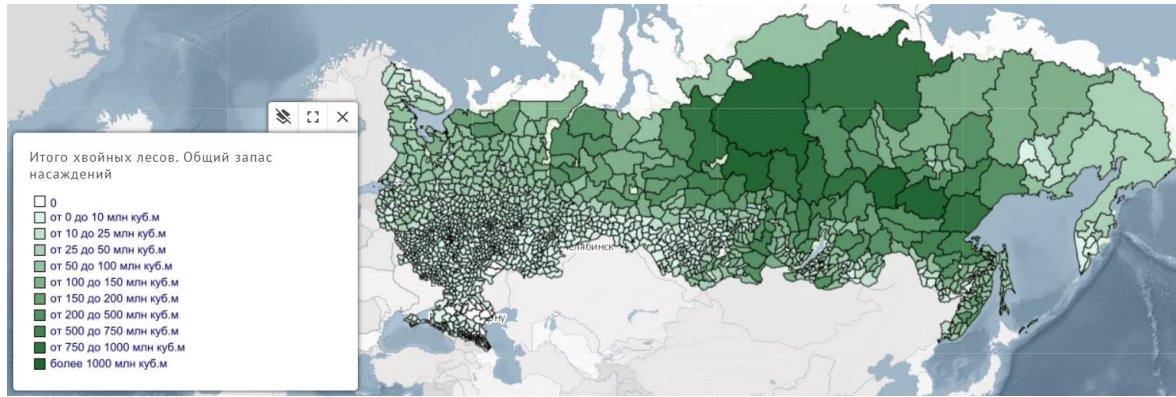


Рисунок 2. Общий запас насаждений хвойных лесов в РФ

Стоимость лесных ресурсов в мире Бостонская консалтинговая группа (BCG) оценивает в \$50–150 трлн, а в России – в \$4-17 трлн. По мнению экспертов компании, основной ценностью лесов является роль в регулировании климата (способность поглощать CO₂ и хранить углерод): в мире на нее приходится 65-90%, в России – 81-96% стоимости. При оценке тонны углекислого газа в \$135 стоимость российских лесов будет уже в 5,5 раза выше (\$17 трлн) стоимости запасов нефти [6]. Поэтому сохранение российских лесов является одной из самых важных задач нашего государства, которой, к сожалению, препятствуют регулярные природные пожары.

Классификация лесных пожаров

В настоящей работе рассматриваются только лесные пожары, классификация которых предложена в [7], изображенная на рисунке 3.

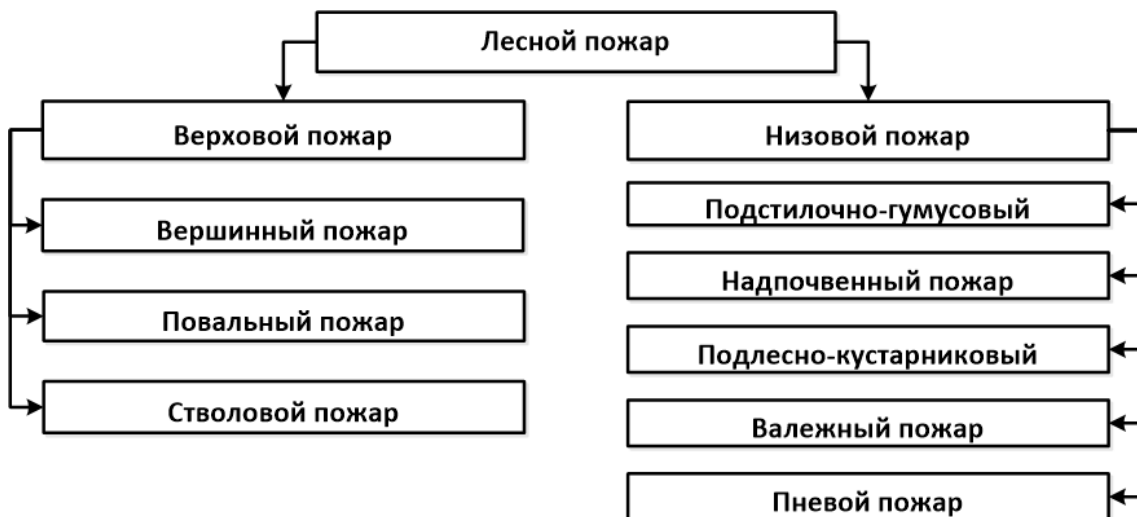


Рисунок 3. Классификация лесных пожаров

Открытые источники информации о лесных пожарах

Основным источником официальной информации о лесных пожарах является ФБУ «Авиалесоохрана», подведомственное Рослесхозу. Это учреждение предоставляет практически информацию, которую далее распространяют другие ведомства и службы.

Существует также система оперативного мониторинга природных пожаров, разработанная российской компанией «СКАНЭКС». Она базируется на технологиях анализа спутниковой информации.

Стоит упомянуть и зарубежные системы. С помощью системы Sentinel Hub возможно создание примерной истории лесных пожаров на определённой территории, демонстрируя визуализации с разных спутников для изучения события. Он включает в себя данные о концентрации газа с Sentinel-5P, тепловизионные снимки с Landsat 8 и различные визуализации пожара и сгоревшей территории с Sentinel-2. Основным и самым удобным источником обзорных спутниковых данных о лесных пожарах является сайт Worldview, созданный и поддерживаемый Национальным управлением по авионавтике и исследованию космического пространства США – NASA.

Инфологическая модель базы данных

На основе анализа информации из рассмотренных выше источников предложена следующая инфологическая структура базы данных о лесных пожарах, представленная на рисунке 4.

База данных позволяет хранить всю необходимую и в то же время достаточную информацию, для реализации технологии прогнозирования распространения лесных пожаров в том числе в реальном времени при наличии технической возможности наблюдения с применением средств аэрофотосъемки.

Алгоритм сбора информации и анализа полученных данных

Для получения информации был проведён анализ веб-страниц сайта ФБУ «Авиалесоохрана». Выполненный анализ показал возможность извлечения фактических данных. На рисунке 5 представлен укрупненный алгоритм сбора данных.

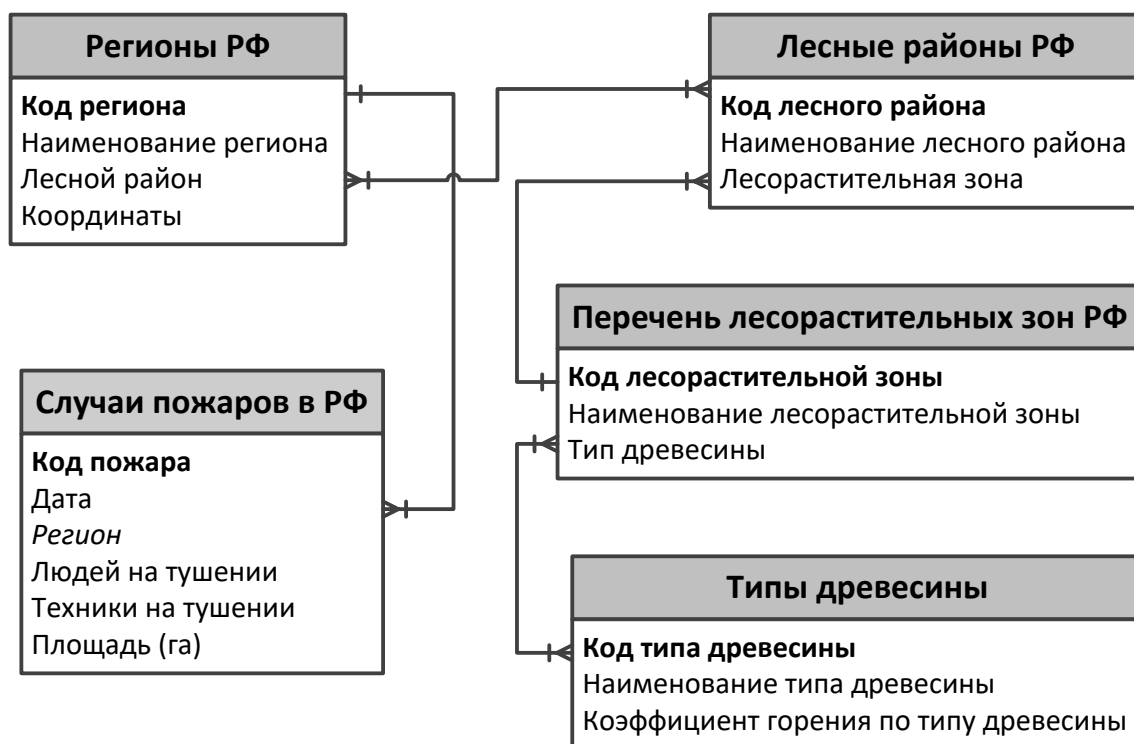


Рисунок 4. Модель данных разработанной информационной системы

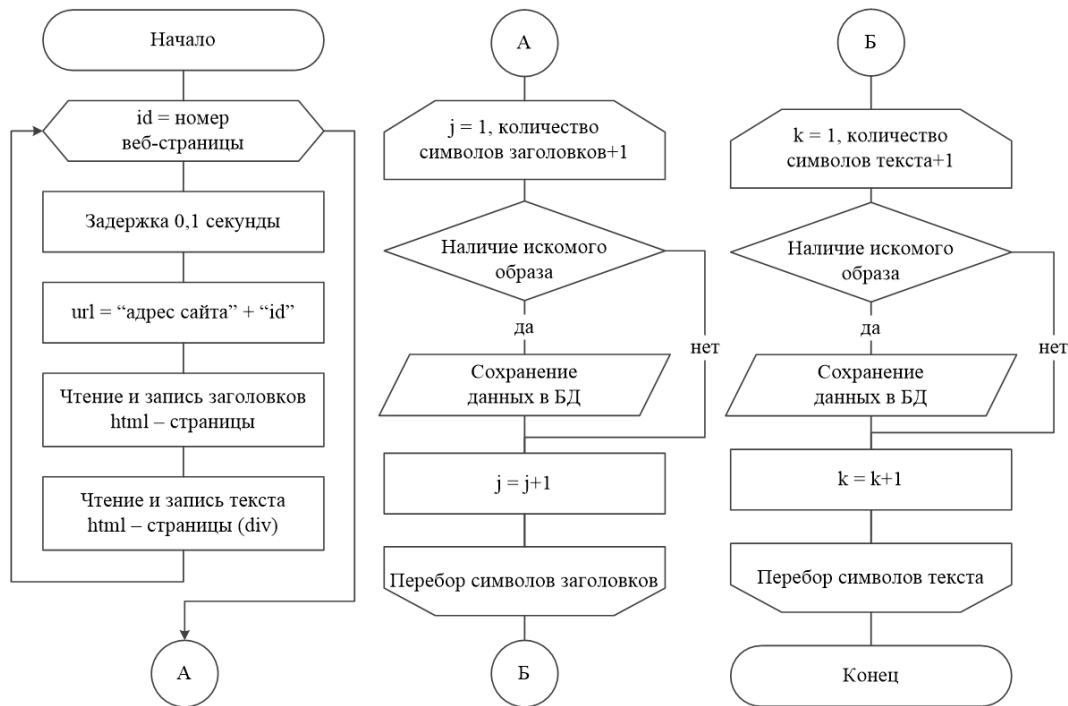


Рисунок 5. Укрупненный алгоритм автоматизированного сбора данных

Выводы

В результате проделанной работы была сформирована база данных о лесных пожарах, произошедших на территории РФ за последние 15 лет. База данных предназначена для оценки корректности работы алгоритмов цифрового двойника лесных пожаров, разрабатываемого в целях прогнозирования развития пожарной ситуации в реальном масштабе времени, предназначенная для систем поддержки принятия решений при парировании нештатных ситуаций.

Разработанные модели и алгоритмы автоматизированного сбора информации позволяют существенно сократить время на сбор и анализ данных значительного объема, анализ которых позволит быстро оценить ситуацию и принять решение о целесообразности продолжения работы в выбранном направлении в любой сфере человеческой деятельности, включая обеспечение безопасности и комфорта населения.

При реализации проекта использовались свободно-распространяемое инструментальное программное обеспечение: Python, СУБД MySQL (Maria DB). Замена инструментов на их отечественные аналоги не представит больших сложностей, так как и то и другое доступно в виде открытого исходного кода.

Литература

1. Лесоводство: учебник/С.В. Залесов. Екатеринбург: УГЛТУ, 2020. – 295 с.
2. Федеральное бюджетное учреждение «Авиалесоохрана». Официальный сайт. URL: <https://aviales.ru> (дата обращения: 24.03.22).
3. Лес. Типы, виды и классификация лесов.//Современные технологии производства. URL:<https://extxe.com/15639/les-tipy-vidy-i-klassifikacija-lesov/> (дата обращения: 07.04.22).
4. Лес. Большая российская энциклопедия – электронная версия. URL: <https://bigenc.ru/biology/text/2141479> (дата обращения: 24.03.22).

5. Система оперативного мониторинга СКАНЭКС, сервис «Карта пожаров». URL: <https://fires.ru/help.html> (дата обращения: 29.03.22).

6. Дзядко Т. Международные эксперты оценили стоимость Российских лесов. URL: <https://www.rbc.ru/business/11/08/2020/5f3172449a794755a58729a5> (дата обращения: 24.03.22).

7. Сафронов, М.А. Ещё раз о классификациях лесных пожаров//Лесн. хоз-во. – 1971. – № 2. – С. 22-25.

УДК 004.75

**ПРОГРАММНОЕ УПРАВЛЕНИЕ КОМПЬЮТЕРНЫМИ УЗЛАМИ
В УСЛОВИЯХ ИЗМЕНЯЮЩИХСЯ ПОТОКОВ ЗАЯВОК
С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА НАЗИНА-ПОЗНЯКА**

**SOFTWARE MANAGEMENT OF COMPUTER NODES
IN CONDITIONS OF CHANGING FLOW
THE NAZIN-POZNYAK ALGORITHM**

Ткаченко К.С.,
ФГАОУ ВО «Севастопольский государственный университет»,
г. Севастополь, Российская Федерация

K.S. Tkachenko,
FSAEI HE “Sevastopol state university”,
Sevastopol, Russian Federation

e-mail: KSTkachenko@sevsu.ru

Аннотация. Для управления компьютерными сетями часто применяются активные анализаторы сетевого трафика. Благодаря ним становится возможным производить выявление недостатков функционирования компьютерных сетей. На их основе можно построить средства управления компьютерной сетью для улучшения качества функционирования. Если системы управления сетями являются интеллектуальными, то возможно принимать нестандартные решение по управления компьютерными сетями в условиях неопределенности и случайности. Для управления в условиях неопределенности и случайности применяются алгоритмы стохастической оптимизации. Эти алгоритмы позволяют адаптивно регулировать параметры компьютерной сети. В работе рассматривается алгоритм Назина-Позняка. Этот алгоритм является проекционным, то есть рекуррентное соотношение для его реализации использует оператор проецирования на симплекс. По результатам анализа выявленных недостатков функционирования компьютерных сетей можно изменить характеристики компьютерных узлов. С использованием алгоритмов стохастической аппроксимации можно произвести прогнозирование работы компьютерной сети во времени и улучшить качество ее функционирования. На языке программирования высокого уровня Java был разработан программный комплекс, который позволяет организовать оптимизацию с использованием различных проекционных и беспроекторных алгоритмов, в том числе, и Назина-Позняка.

Abstract. Active network traffic analyzers are often used to manage computer networks. Thanks to them, it becomes possible to identify the shortcomings of the functioning of computer networks. On their basis, it is possible to build computer network management tools to improve the quality of functioning. If network management systems are intelligent, it is possible to make non-standard decisions on computer network management in conditions of uncertainty and randomness. Stochastic optimization algorithms are used for control under conditions of uncertainty and randomness. These algorithms allow you to adaptively adjust the parameters of a computer network. The paper considers the Nazin-Poznyak algorithm. This algorithm is a projection algorithm, that is, the recurrence relation for its implementation uses the projection operator on the simplex. Based on the results of the analysis of the identified shortcomings in the functioning of computer networks, it is possible to change the characteristics of computer nodes. Using stochastic approximation algorithms, it is possible to predict the operation of a computer network over time and improve the quality of its functioning. A software package has been developed in the high-level programming language Java, which allows you to organize optimization using various projection and non-projection algorithms, including Nazin-Poznyak.

Ключевые слова: компьютерные узлы, поддержка принятия решений, стохастическая аппроксимация, проекционные алгоритмы, программная система, управление.

Keywords: computer nodes, decision support, stochastic approximation, projection algorithms, software system, management.

Для управления компьютерными сетями часто применяются активные анализаторы сетевого трафика [1]. Эти анализаторы также позволяют производить управление различными удаленными компьютерными системами по результатам их мониторинга. Аппаратное и программное обеспечение таких анализаторов должно быть согласовано с аналогичным обеспечением мониторируемых компьютерных сетей. Но, в отличие от эксплуатируемых в компьютерной сети компьютерных узлов, в процессе работы анализатора им производится обработка не только адресованных ему информационных потоков, а всех циркулирующих в сети пакетов данных. Традиционно анализаторы имеют сложную структуру, включающую в свой состав сетевые средства, функционирующие на различных уровнях сетевых протоколов. Мониторинг может быть улучшен при наличии в составе анализатора экспертных систем либо их отдельных компонентов и подсистем. Благодаря наличию таких подсистем становится возможным производить выявление таких недостатков функционирования компьютерных сетей, которые других образом определить невозможно. По результатам мониторинга формируются оценочные показатели сетевого функционирования. Также внедрение специализированных агентов приводит к возможности получения к последующей обработке пакетов данных из недоступных либо изолированных внутренних сетевых сегментов.

Рассмотрение и описание компьютерной сети как функционального объединения компьютерных элементов позволяет производить оптимизацию сетевых рабочих процессов [2]. В качестве сетевых характеристик, над которыми возможно проведение процессов оптимизации, часто используются скорость передачи данных, динамическое изменение перегрузок, динамическое изменение загруженности буферов. Качество функционирования компьютерной сети может быть описано на основе этих характеристик. Для улучшения качества функционирования можно использовать средства управления компьютерной сетью. Эти средства управления ложатся в основу

прогнозирования изменений сетевых характеристик, что позволяет адаптивно регулировать параметры компьютерной сети. Системы управления сетью могут являться интеллектуальными, интеллектуальный характер работы систем приводит к возможностям принятия нестандартных решений по управления компьютерными сетями, а также к допустимости их функционирования в условиях неопределенности и случайности. Это означает, что с их применением становится возможным формировать принципиально новые алгоритмы функционирования компьютерных сетей. Интеллектуальные системы управления должны работать под различными операционными системами и аппаратными окружениями, как на отдельных компьютерных узлах, так и в специально выделенных системах, с участием системных администраторов и без них.

Поэтому далее в настоящей работе рассматривается программное управление компьютерными узлами в условиях изменяющихся поток заявок с использованием алгоритма Назина-Позняка.

Алгоритм Назина-Позняка является одним из наиболее важных алгоритмов принятия решений в условиях неопределенности [3]. Это позволяет производить построение на его основе программных систем поддержки принятия решений [4-6].

Пусть мониторируемый компьютерный узел, подверженный изменяющемуся во времени потоку заявок, обладает конечным множеством управляющих воздействий $x(1), \dots, x(N)$. Тогда для управления ним можно использовать алгоритмы стохастической аппроксимации [3]. Существуют различные алгоритмы стохастической аппроксимации, к их числу относят алгоритмы Нарендры-Шапиро, Льюса, Варшавского-Воронцовой, Буша-Мостеллера, Назина-Позняка и некоторые другие.

В общем случае для программного управления на основе алгоритмов стохастической оптимизации можно использовать рандомизированные стратегии. В рандомизированных стратегиях используются рекуррентные правила вида:

$$p_{n+1} = R_n(x_1, \dots, x_n; p_1, \dots, p_n; \xi_1, \dots, \xi_n), \quad n = 1, 2, \dots \quad (1)$$

В формуле (1) R_n – вектор-функция со значениями в симплексе S^N , p_n – вектор условных вероятностей выбора вариантов $x(1), \dots, x(N)$ в момент времени t_n , $n = 1, 2, \dots$ – номер шага, ξ_n – текущие потери. Перед выбором очередного варианта x_{n+1} происходит расчет p_{n+1} по (1). Выбор варианта осуществляется методом деления отрезка. Для функционирования в условиях неограниченной штрафной функции необходимо ввести оператор проектирования на симплекс. Пусть для любого $q \in R^N$ существует вектор-столбец $p = \pi_\varepsilon^N\{q\}$, принадлежащий $(N-1)$ -мерному единичному ε -симплексу ($\varepsilon_n \in [0, N^{-1})$ – параметр проектора):

$$S_\varepsilon^N = \left\{ p = (p_1, \dots, p_N) \mid p \in R^N, \sum_{i=1}^N p_i = 1, p_i \geq \varepsilon \ (i = 1, \dots, N) \right\} \quad (2)$$

В формуле (2) вектор-столбец

$$p = \pi_\varepsilon^N\{q\}$$

рассчитывается для обеспечения выполнения условия

$$\|\pi_{\varepsilon}^N\{q\} - q\| = \min_{p \in S_{\varepsilon}^N} \|p - q\|.$$

Для любого $q \in R^N$ вектор $\pi_{\varepsilon}^N\{q\}$ должен быть единственным. Существующие и применяемые на практике алгоритмы адаптивного выбора вариантов группируются на беспроекторные алгоритмы адаптивного выбора вариантов вида

$$p_{n+1} = p_n - \gamma_n R(x_n, p_n, \xi_n)$$

и проекционные алгоритмы адаптивного выбора вариантов вида

$$p_{n+1} = \pi_{\varepsilon_{n+1}}^N \{p_n - \gamma_n R(x_n, p_n, \xi_n)\},$$

где $R(x_n, p_n, \xi_n) = R_n$ – вектор движения алгоритма,

$\gamma_n > 0$ – скалярный множитель – длина шага.

К числу проекционных алгоритмов относится и алгоритм Назина-Позняка.

Алгоритм Назина-Позняка определяется рекуррентным соотношением вида:

$$p_{n+1} = \pi_{\varepsilon_{n+1}}^N \left\{ p_n - \gamma_n \frac{\xi_n - \Delta}{e^T(x_n) p_n} e(x_n) \right\}. \quad (3)$$

В формуле (3) Δ – декремент для нормализации.

На языке программирования высокого уровня Java был разработан программный комплекс, который позволяет организовать оптимизацию с использованием различных проекционных и беспроекторных алгоритмов, в том числе, и Назина-Позняка [4]. Результаты работы приводятся на рисунках 1 и 2.

К.С. Ткаченко. Инструментальные средства информационного обеспечения принятия решений в распределенных средах

Алгоритм Буша-Мостеллера	Без адаптации	Алгоритм 12VRS	Робастный	Результаты работы	График функции текущих средних потерь
Проекционный алгоритм стохастической аппроксимации по Назину-Позняку		Алгоритм Нарендры-Шапиро		Алгоритм Льюса	Алгоритм Варшавского-Воронцовой
Число вариантов					5
Число шагов					50
delta					0,5
gamma начальное					0,5
epsilon начальное					0,1
hi					0,5
pi					0,4
A					0,1
B					0,1

Теоретические потери

Расчет

Рисунок 1. Параметры алгоритма Назина-Позняка

К.С. Ткаченко. Инструментальные средства информационного обеспечения принятия решений в распределенных средах

Алгоритм Буша-Мостеллера	Без адаптации	Алгоритм 12VRS	Робастный	Результаты работы	График функции текущих средних потерь
Проекционный алгоритм стохастической аппроксимации по Назину-Позняку			Алгоритм Нарендры-Шапиро	Алгоритм Льюиса	Алгоритм Варшавского-Воронцовой

График величины текущих средних потерь

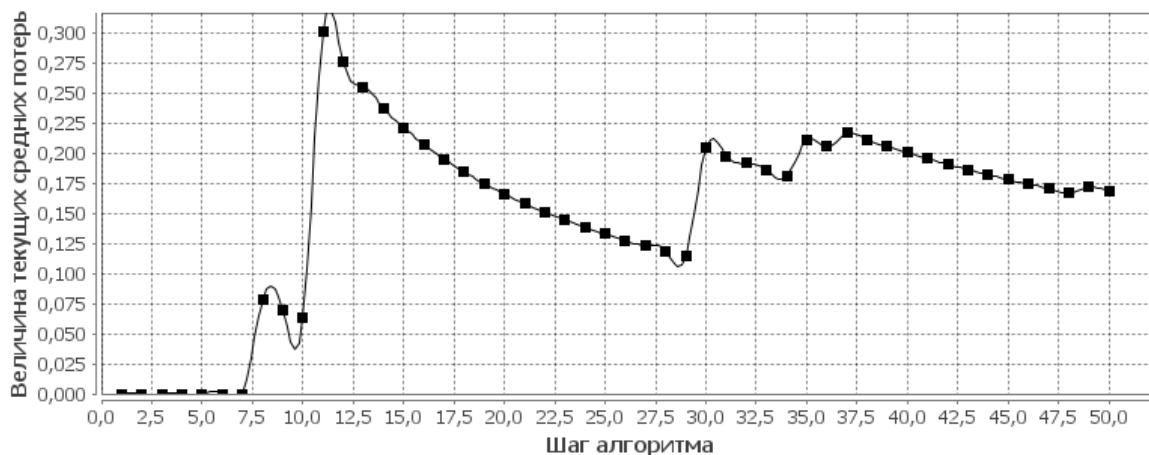


Рисунок 2. График величины текущих средних потерь

Применение программного комплекса для реализации оптимизации на основе алгоритма Назина-Позняка приводит к уменьшению модельных предельных значений средних текущих потерь.

Выводы

Для сетевых компьютерных узлов можно формировать управляющие воздействия с использованием алгоритмов стохастической аппроксимации. Разработанный программный комплекс позволяет производить параметрическую коррекцию компьютерного узла для уменьшения модельных предельных значений средних текущих потерь.

Литература

1. Муминов Б.Б., Тургунов Ш.Ш. Анализ протоколов управления компьютерными сетями//Science and Education, т.3, №1, 2022. С. 246-254.
2. Леохин Ю.Л., Бондаренко А.Д. Проектирование интеллектуальных систем управления компьютерными сетями//Лесной вестник, №2, 2007. С. 149-152.
3. Назин А.В., Позняк А.С. Адаптивный выбор вариантов: Рекуррентные алгоритмы. М.: Наука, 1986. 288 с.
4. Ткаченко К.С. Программный комплекс информационного обеспечения принятия решений в распределенных средах//Свидетельство о регистрации программы для ЭВМ 2021617213, 13.05.2021. Заявка № 2021613477 от 16.03.2021.
5. Ткаченко К.С. Программная система для выбора вариантов на основе проекционных и беспроекционных алгоритмов//Системы контроля окружающей среды. 2016. №3(23). С. 59-62.
6. Ткаченко К.С. Совершенствование средств компьютерной безопасности в организациях путем проведения узловой параметрической корректировки//Вестник Прикамского социального института. 2021. № 2 (89). С. 87-92.

УДК 004.415

ДОБАВЛЕНИЕ ОБЪЕКТА В ГЕОСОЦИАЛЬНОЙ СЕТИ «MYTRIP»**ADDITION OF AN OBJECT IN GEOSOCIAL NETWORK «MYTRIP»**

Овсянникова А.С., Белозеров А.Е.,
Уфимский государственный нефтяной технический университет,
ул. Космонавтов, 1, г. Уфа, Республика Башкортостан, 450064, Россия

A.S. Ovsyannikova, A.E. Belozerov,
Ufa State Petroleum Technological University,
Kosmonavtov Str., 1, Ufa, Republic of Bashkortostan, 450064, Russia

e-mail: anna.s.ovsyannikova@mail.ru

Аннотация. Наш мир наполнен удивительными местами – как рукотворными, так и созданными природой. Существует множество локаций и достопримечательностей, которые ежегодно привлекают к себе внимание тысяч туристов. В настоящее время есть множество информационных ресурсов, которые помогают получить нужную информацию об интересующем месте или объекте. К таким сервисам относятся: справочные системы, картографические сервисы и сервисы для путешествий, а также различные интернет-ресурсы, социальные и геосоциальные сети. Геосоциальная сеть «MyTrip» объединяет знания пользователей со всего мира об интересных местах и объектах в одном приложении и рекомендует место проведения досуга по предпочтениям пользователя. В данной статье описано добавление нового объекта в геосоциальную сеть «MyTrip». Проведен анализ существующих на данный момент подходов к обработке и накоплению данных, определены их достоинства и недостатки. Изучены виды модерации. Приведено решение проблемы появления недостоверной информации. Рассмотрены особенности добавления объекта в существующее и в новое здание. Описан подход к сбору основной и дополнительной информации об объекте. Разработанный алгоритм добавления новых объектов используется при создании геосоциальной сети и, в частности, мобильного приложения.

Abstract. Our world is filled with amazing places – both created by nature or human. There are many locations that attract the attention of thousands of tourists every year. Currently, there are many information resources that help to get the necessary information about the place or object. Such services include: reference systems, mapping and travel services, different Internet resources, social and geosocial networks. Geosocial network “MyTrip” combines the knowledge of users from all over the world about interesting places and objects in one application and recommends a place according to the user's preferences.

This article describes the addition of a new object in geosocial network “MyTrip”. The analysis of the currently existing approaches to data processing and accumulation is carried out, their advantages and disadvantages are determined. The types of moderation have been studied. The solution of the problem of the appearance of false information is given. The features of adding an object to an existing building and to a new building are considered. The approach to collecting basic and additional information about the object is described. The developed algorithm for adding new objects is used for creating a geosocial network and, in particular, a mobile application.

Ключевые слова: объект, геосоциальная сеть, накопление данных, местоположение, информационная система, мобильное приложение.

Keywords: object, geosocial network, data accumulation, location, information system, mobile application.

В мире существует множество удивительных мест, созданных природой или человеком. Среди них есть как очень популярные, так и вовсе не известные. В век технологий доступной информации становится больше, и найти интересное место или заведение становится сложнее. Главными помощниками в данном вопросе становятся друзья и знакомые, которые делятся своим опытом и впечатлениями; специализированные сервисы; социальные сети, веб-сайты.

В настоящее время существует два подхода к обработке и накоплению данных: централизованный и децентрализованный.

При централизованном подходе информация заполняется сотрудниками сервиса. Сотрудники обзванивают организации, самостоятельно делают фотографии, встречаются с владельцами заведений и накапливают базу данных. Данный метод хорош тем, что в системе отсутствует спам. Однако для наполнения новой информацией и обновления существующей, требуется много времени, большой штат сотрудников и, следовательно, средств для оплаты их труда.

Децентрализованный подход позволяет пользователю самому добавлять информацию, которая кажется ему интересной и актуальной. Пользователи сервиса при посещении какого-либо объекта, добавляют его в систему, делают фотографии, добавляют отзывы. В данном случае информация, хранящаяся в системе, быстро пополняется и актуализируется. Недостатком данного подхода является то, что в системе возможно появление недостоверной информации.

Многие информационные системы используют механизмы модерации. Существует два вида модерации: премодерация и постмодерация.

Пользователь заполняет всю необходимую информацию и отправляет ее на премодерацию. После проверки материала на соответствие правилам модератор либо отправляет информацию на редактирование, либо размещает данные в общий доступ. Премодерация затрудняет добавление новой информации в систему и может вызвать потерю интереса у пользователей.

При постмодерации информация проверяется уже после публикации, и пользователь может сразу увидеть добавленный объект.

В геосоциальной сети «MyTrip» используется децентрализованное совместное накопление информации в структурированном виде, так как актуальность представленных данных является главной составляющей проекта. Данные предоставляются пользователям на основе рейтинга [1]. Для решения проблем дублирования и недостоверности информации применяется постмодерация.

Для работы с геосоциальной сетью разрабатывается мобильное приложение. Пользователи могут получить доступ к имеющейся информации, а также добавлять новые объекты в любое время и в любом месте.

Объект – это здание или заведение, имеющее юридического владельца и реализующее экономическую деятельность. Например, кинотеатр, ресторан, торговый центр, магазин.

Объект в систему может добавить любой авторизованный пользователь. На рисунке 1 представлена диаграмма вариантов использования, отображающая взаимодействие пользователя с геосоциальной сетью при добавлении нового объекта.

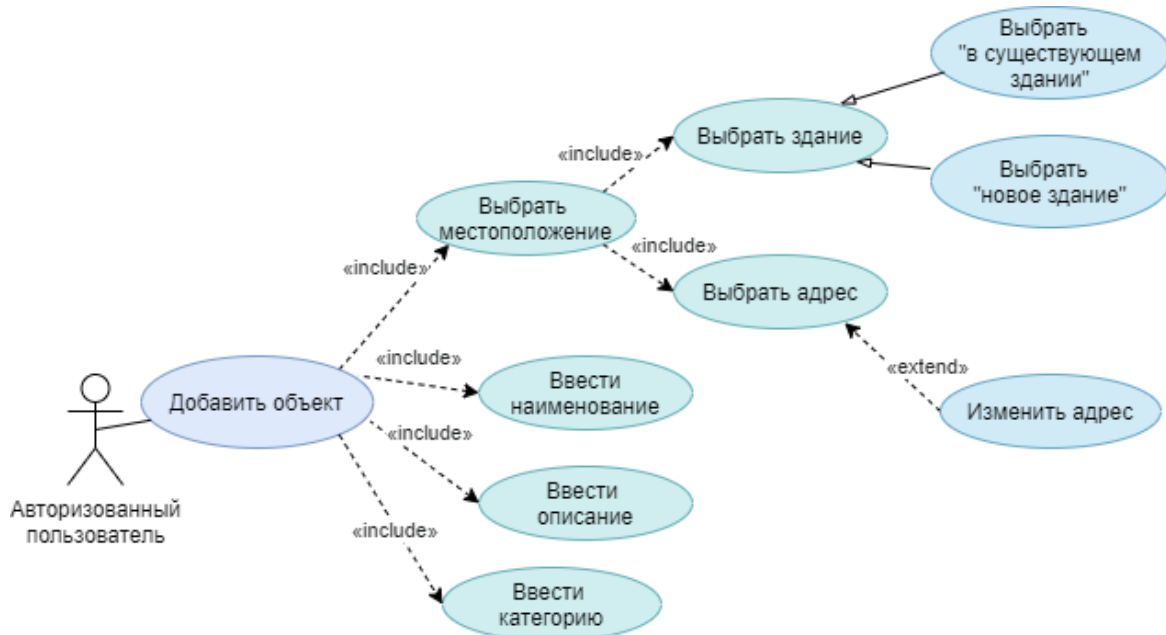


Рисунок 1. Добавление объекта

Для того чтобы добавить новый объект, необходимо выбрать его местоположение на карте. Добавить объект можно либо в уже существующее здание, либо в новое здание, которого еще нет на карте. В первом случае при выборе местоположения объекта маркер всегда перемещается к ближайшему существующему зданию (рисунок 2). Во втором случае присваиваются заданные координаты, и маркер не притягивается к зданиям (рисунок 3).

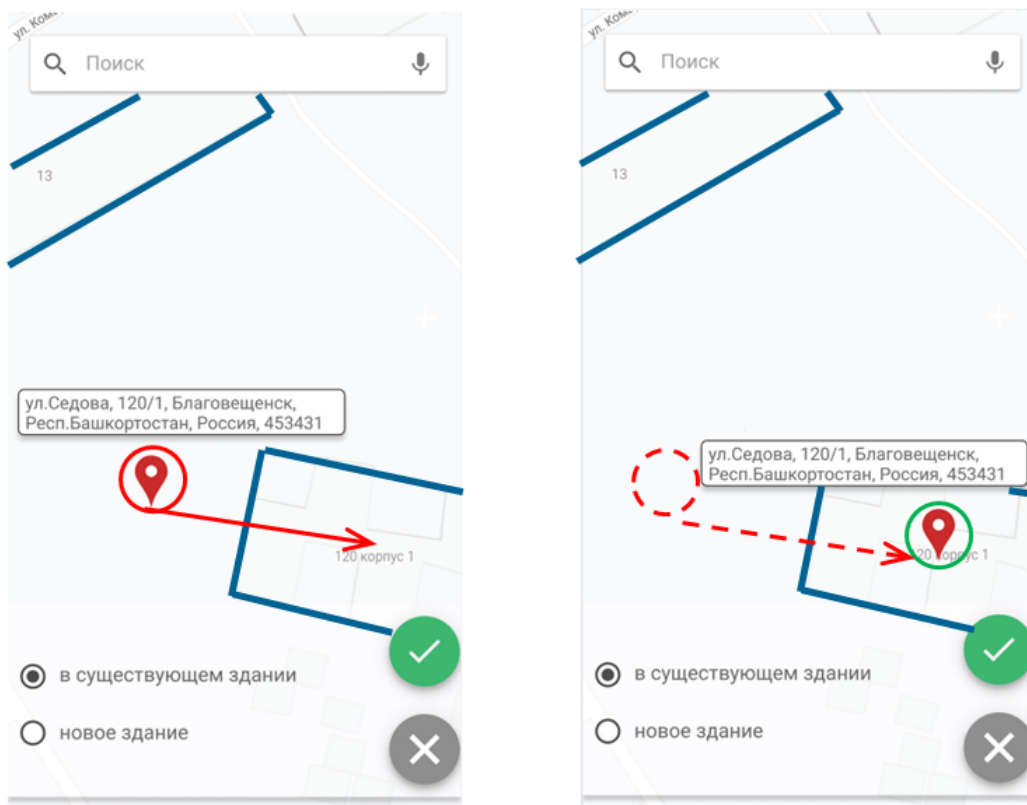


Рисунок 2. Добавление объекта в существующее здание

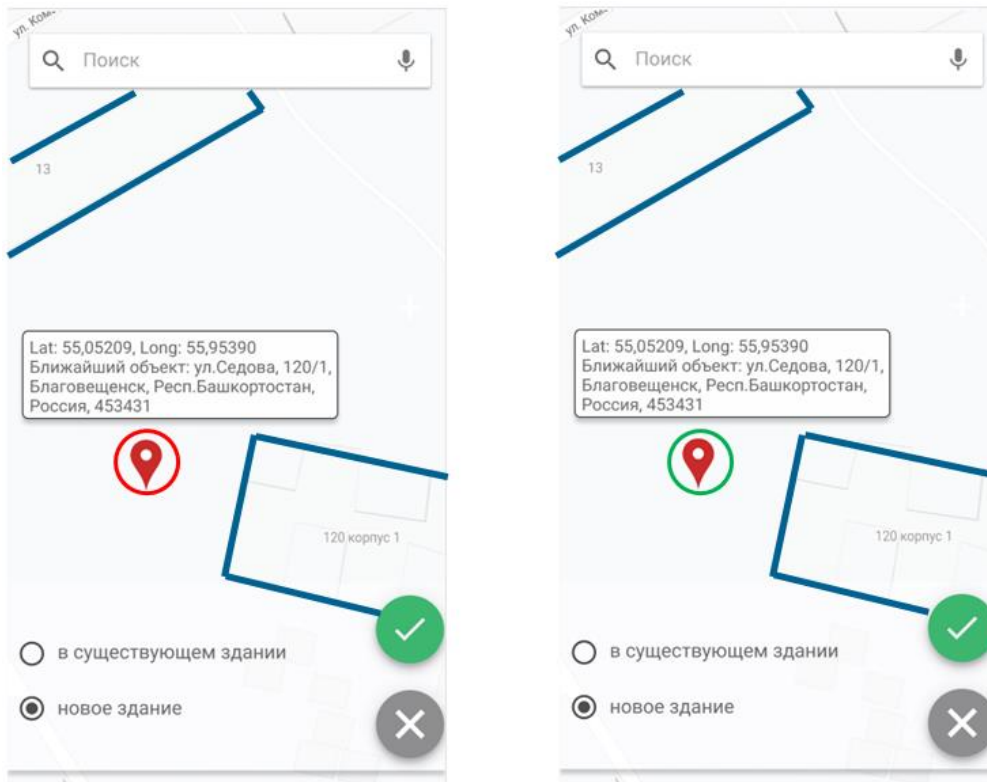


Рисунок 3. Добавление объекта в новое здание

После выбора местоположения необходимо заполнить обязательные поля: наименование, описание и категория – это основная информация об объекте (рисунок 4).

Рисунок 4. Основная информация об объекте

Эмпирическое правило в дизайне формы гласит, что чем она короче, тем лучше. Меньше усилий со стороны пользователя приведет к повышению конверсии [2]. Таким образом, задача состояла в том, чтобы свести к минимуму количество обязательных

полей. В противном случае, пользователь может утратить интерес из-за необходимости заполнения большого количества полей одновременно.

Все объекты, созданные пользователем, отображаются в личном кабинете (рисунок 5) во вкладке «Мои объекты».

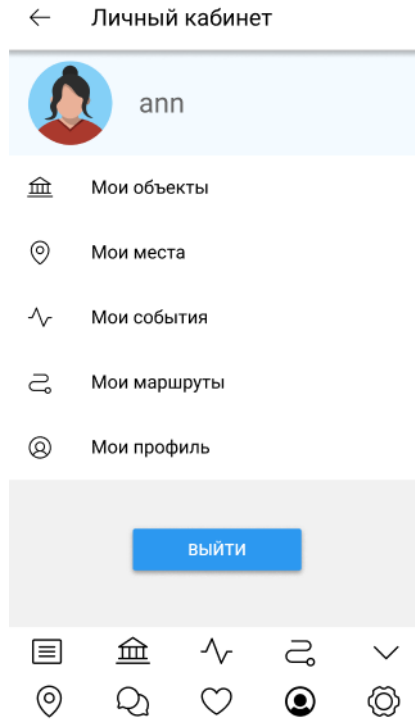


Рисунок 5. Макет личного кабинета пользователя

Личный кабинет позволяет управлять созданными объектами. После добавления объекта здесь можно внести дополнительную информацию, заполнение которой и было опущено ранее с целью экономии времени пользователя (рисунок 6). К полям, не обязательным для заполнения, относятся: описание, средний чек, контактная информация (телефон, сайт и социальные сети) и возрастное ограничение.

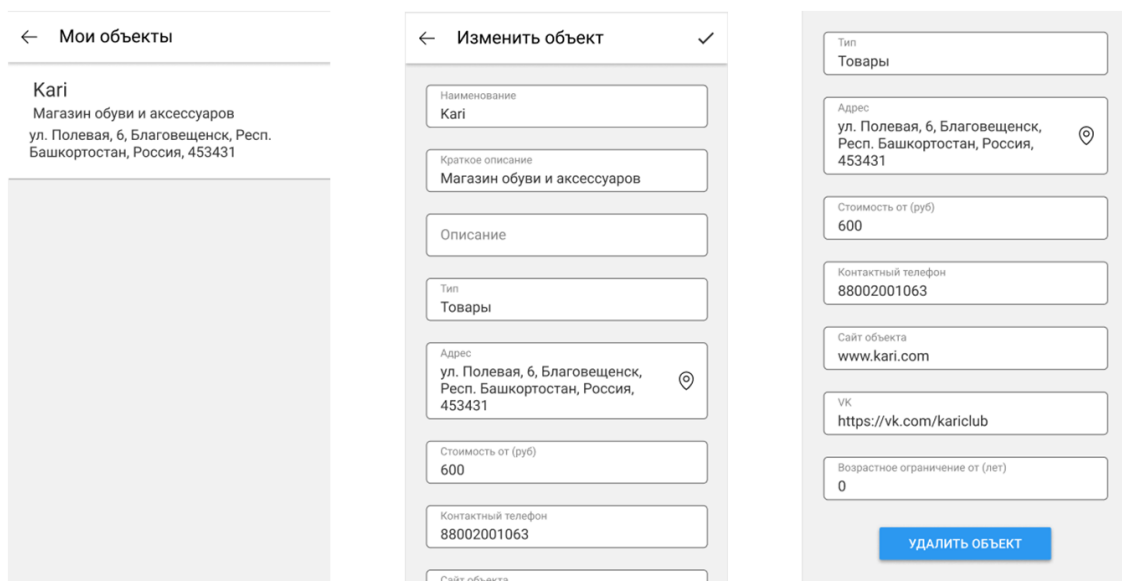


Рисунок 6. Заполнение полей с дополнительной информацией об объекте

Выводы

В рамках создания геосоциальной сети «MyTrip» разрабатываются механизмы и функционал добавления нового объекта в систему. Решены задачи выбора местоположения объекта по отношению к другим зданиям, ввода основной и дополнительной информации об объекте. Разработан просмотр и редактирование объектов через личный кабинет. Реализован общий доступ к объектам через геосоциальную сеть.

Литература

1. Фукалов А.С., Белозеров А.Е. Расчет рейтинга мероприятия в геосоциальной сети. – Информационные технологии. Проблемы и решения. – Уфа: Изд-во УГНТУ, 2020. № 4 (13). С. 51-54.
2. Designing More Efficient Forms: Structure, Inputs, Labels and Actions [Электронный ресурс]. – URL:<https://medium.com/thinking-design/designing-more-efficient-forms-structure-inputs-labels-and-actions-7781d9b6d32c> (Дата обращения: 24.04.2022).

СЕТИ И ТЕЛЕКОММУНИКАЦИИ

УДК 004.72

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ АНАЛИЗА РАЗМЕЩЕНИЯ АППАРАТНО-ПРОГРАММНЫХ КОМПОНЕНТОВ В АРХИТЕКТУРЕ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ

DEVELOPMENT OF AN AUTOMATED SYSTEM FOR ANALYSIS OF THE PLACEMENT OF HARDWARE AND SOFTWARE COMPONENTS IN ARCHITECTURE

Гаврилов И.С., Тутубалин П.И.,
Казанский национальный исследовательский технический университет
им. А.Н. Туполева
г. Казань, Российская Федерация

I.S. Gavrilov, P.I. Tutubalin,
Kazan National Research Technical University
named after A.N. Tupolev,
Kazan, Russian Federation

e-mail: ivan_gavrilov_2013@mail.ru

Аннотация. Актуальность темы исследования. При разработке новой сетевой архитектуры или внедрении какого-либо компонента в существующую проводится этап проектирования, одним из результатов которого является схема целевой сети. На данный момент для её отображения используют специальное программное обеспечение построения диаграмм. Оно обрело свою популярность благодаря тому, что предоставляет большой и удобный инструментарий для изображения различного вида схем, которые на выходе имеют удобный для восприятия человеком вид. Кроме того, формат результирующего файла построенной диаграммы позволяет без затруднений производить корректировки, конвертировать в файлы других типов, делиться им и включать в электронную версию документации. Целью исследования является изучение автоматизированного подхода к анализу размещения аппаратно-программных компонентов в архитектуре вычислительной сети с учётом обеспечения безопасности. В статье описана структура файла наиболее популярного формата изображения диаграмм, раскрыта идея о последующих преобразованиях, содержащихся в нём данных, приведено обоснование использования формального языка описания правил проверки. Была разработана процедура автоматизированного анализа схем вычислительных сетей. В результате исследования были выделены компоненты системы, описаны принципы её функционирования и возможная архитектура.

Abstract. Relevance of the research topic. When developing a new network architecture or introducing a component into an existing one, a design stage is carried out, one of the results of which is a diagram of the target network. At the moment, special diagramming software is used to display it. It has gained its popularity due to the fact that it provides a large and convenient toolkit for depicting various types of circuits, which at the output have a human-friendly appearance. In addition, the format of the resulting file of the constructed diagram

allows you to easily make adjustments, convert to files of other types, share it and include it in the electronic version of the documentation. The aim of the study is to study an automated approach to the analysis of the placement of hardware and software components in the architecture of a computer network, taking into account security. The article describes the file structure of the most popular diagram image format, reveals the idea of subsequent transformations of the data contained in it, and provides a rationale for using a formal language for describing the rules for checking. A procedure for automated analysis of computer network schemes was developed. As a result of the study, the components of the system were identified, the principles of its functioning and possible architecture were described.

Ключевые слова: вычислительная сеть, сетевые устройства, архитектура, проектирование, информационная безопасность.

Keywords: computer network, network devices, architecture, design, cybersecurity.

Введение

На этапе проектирования сетевой архитектуры важно продумать все детали при размещении элементов, так как в дальнейшем допущенная ошибка может стать причиной некорректной работы системы или создать критическую уязвимость информационной безопасности [1-3]. В таких случаях потребуются реорганизация структуры сети системы, что повлечёт за собой дополнительные расходы.

Каждый элемент сети уникален и имеет определенный перечень требований по позиционированию его в сети. И потому метод ручной проверки целевой схемы сети имеет ряд недостатков:

- проверяющий может забыть какую-либо особенность продукта и, в соответствии с этим, не выполнить необходимую проверку
- поиск ошибок и недочётов может занимать довольно много времени
- при отсутствии сотрудника с необходимой компетенции невозможно провести анализ схемы сети или сделать это качественно

Автоматизированная система лишена данных недостатков, т. к. позволяет в кратчайшие сроки находить ошибки в предлагаемой схеме сети и ускоряет процесс построения инфраструктуры.

При исследовании открытых источников упоминаний систем, решающих данную задачу, найдено не было. В связи с чем предлагается следующий подход для её решения.

Структура файлов Microsoft Visio

Специалисты используют различные программы построения диаграмм для изображения целевой схемы сети. Один из самых популярных инструментов для решения данной задачи – *Microsoft Visio*. Для демонстрации процедуры анализа рассмотрим схему сети, построенную в *Microsoft Visio* (

Рисунок 1).

Диаграммы *Microsoft Visio* имеют расширение *.vsdx* и представляют собой контейнер для файлов, который имеет следующую структуру (Рис. 2):

Таким образом, файлы *.vsdx* можно открывать, читать, обновлять, изменять и создавать без автоматизации приложения *Microsoft Visio* [4].

Вся информация о диаграммах расположена в XML-документах. XML (eXtensible Markup Language) – это расширяемый язык разметки, который позволяет в человеко-читаемом формате описывать сущности.

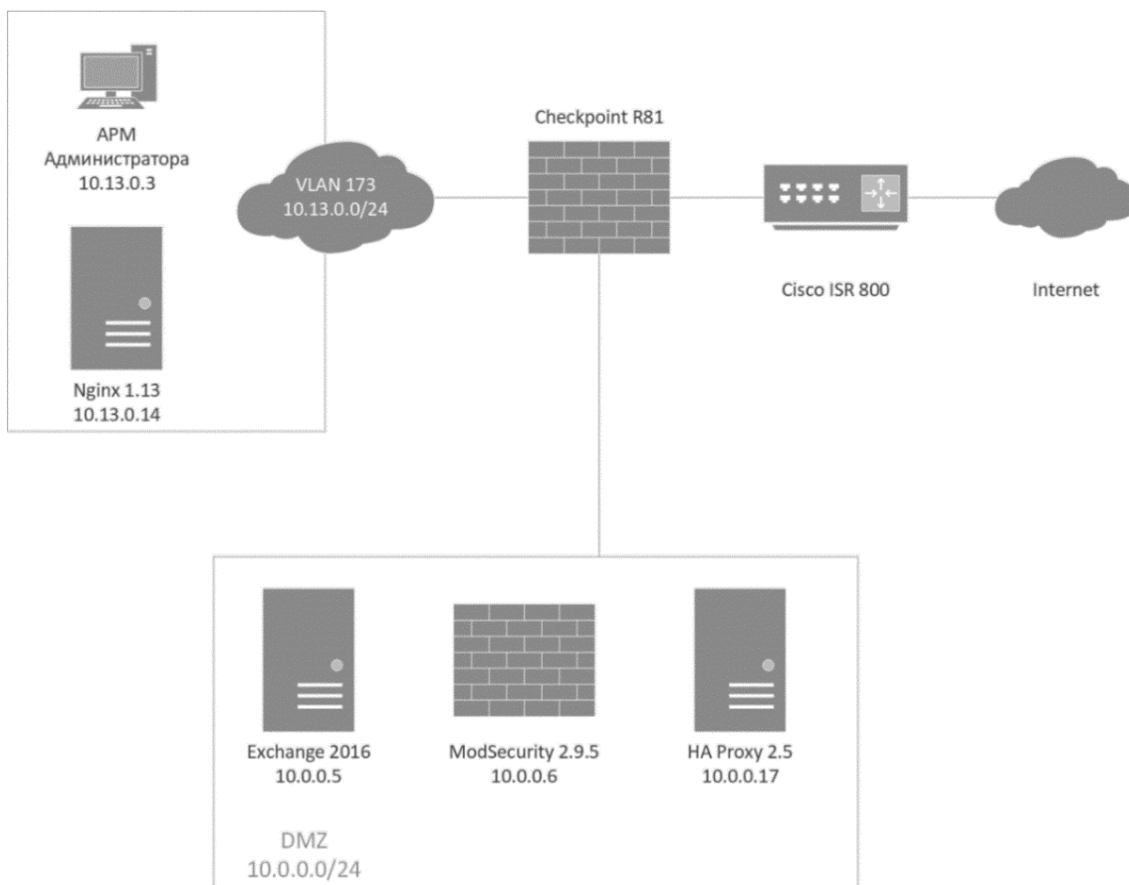


Рисунок 1. Схема сети

Пример описания сущности «Компьютер» в XML-документе:

```
<?xml version="1.1" encoding="UTF-8" ?>
<computer name="pc-1">
  <cpu vendor_name="Intel" cores="4" />
  <ram type="DDR4" value="8GB"/>
  <memory type="ssd" value="500GB">
</computer>
```

Файлы *Microsoft Visio* могут содержать в себе несколько страниц с диаграммами. В папке *visio/pages* хранятся XML-документы *pageN.xml*, которые описывают элементы соответствующей страницы с номером *N*. Файл *visio/masters/masters.xml* содержит перечень всех типов фигур, используемых для построения текущей диаграммы.

Описание XML-сущности фигуры может меняться в зависимости от её типа, но в общем случае оно содержит:

- идентификатор,
- идентификатор типа, описанного в *visio/masters/masters.xml*
- координаты местоположения на диаграмме,
- размер (ширина и высота),
- текст фигуры, его координаты и размер,

Таким образом, логическую структуру XML-документов файла *.vsdx* можно представить в следующем виде (Рисунок 3):

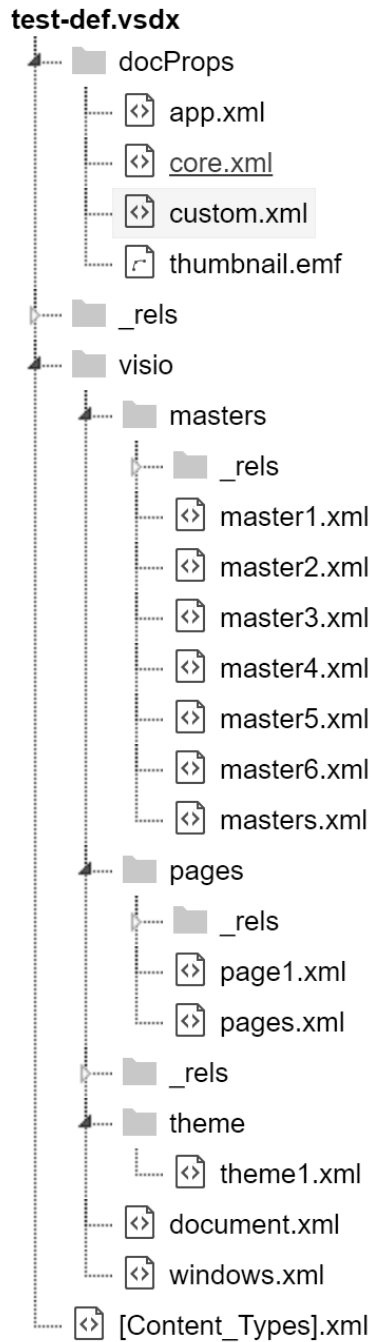


Рисунок 2. Пример содержимого .vsdx файла

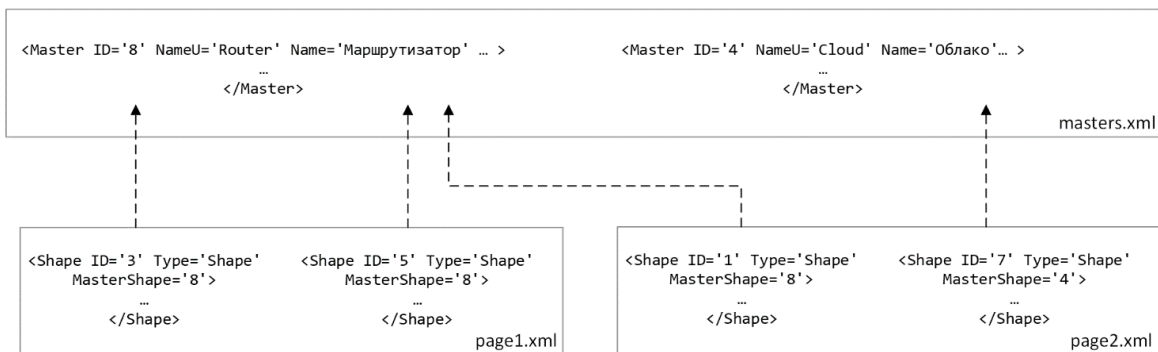


Рисунок 3. Логическая структура XML-документов Microsoft Visio

Описание процедуры анализа схемы

Для выполнения анализа диаграммы необходимо преобразовать содержимое файлов *pageN.xml* в программные сущности и представить их в виде структуры данных. За реализацию данной функции отвечает специальный компонент системы – парсер диаграмм. Выбор структуры зависит от метода анализа схемы, в качестве примера рассмотрим организацию элементов в виде дерева.

В качестве узлов дерева будут выступать отдельные элементы сети и её сегменты. Сегментация сети — это архитектурный подход, который делит сеть на несколько сегментов или подсетей, каждый из которых действует как отдельная небольшая сеть [5]. Связность определяется при помощи соединительных линий на диаграмме – имея координаты их начала и конца (из *XML*-сущности фигуры типа «Динамическая соединительная линия»), а также координаты фигур, можно спроецировать связь на структуру данных. Корневым элементом в дереве будет выступать Интернет, так как с него начинается путь трафика во внутреннюю сеть. Дальнейшие элементы располагаются в зависимости от устройств и сегментов, через которые им необходимо пройти для доступа во внешнюю сеть. Представленная выше схема (Рис. 1) отображается в следующее дерево ():

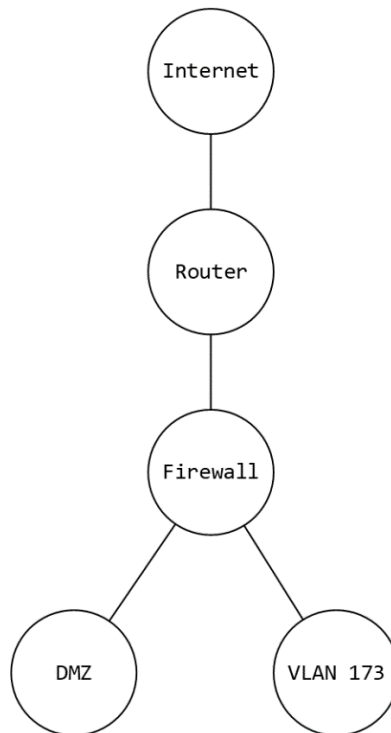


Рисунок 1. Представление элементов сети в виде дерева

В данном случае узлы дерева «DMZ» и «VLAN» содержат в себе информацию о расположенных в соответствующих сегментах сети элементах.

Затем построенное дерево передается в следующий компонент системы - движок проверки. Он, в свою очередь, получает набор правил (например, из базы данных правил) и проверяет их выполнимость на полученной структуре данных. Правила анализа схемы не следуют жестко задавать в программном коде системы, имеет смысл использовать для их описания формальный язык (разработанный язык для решения конкретных проблем [6]), что даёт возможность добавления новых проверок и расширения функционала.

Каждое правило может содержать в себе название и версию элемента сети, к которому оно будет применяться.

Пример описание правила под названием «*nginx_is_protected*» для компонента сети *Nginx* версии 1.13:

```
name: nginx_is_protected
selector:
  element_name: NGINX
  element_version: 1.13
rule: |
element.subnet.type == SubnetType::LAN
AND
element.behind_type(ElementType::WAF)
```

Данное правило говорит о том, что элемент располагается во внутреннем сегменте сети и трафик из/в Интернет до него проксируется через *WAF (Web Application Firewall)* – программные или аппаратно-программные комплексы, предназначенные для обеспечения защиты веб-приложений и различных веб-сервисов [7]).

При неудовлетворении схемы хотя бы одному правилу, система помечает результат проверки как неудачный. Так как имеется возможность модификации диаграмм через *XML*-документы, некорректные моменты могут быть указаны прямо на схеме (рис. 5).

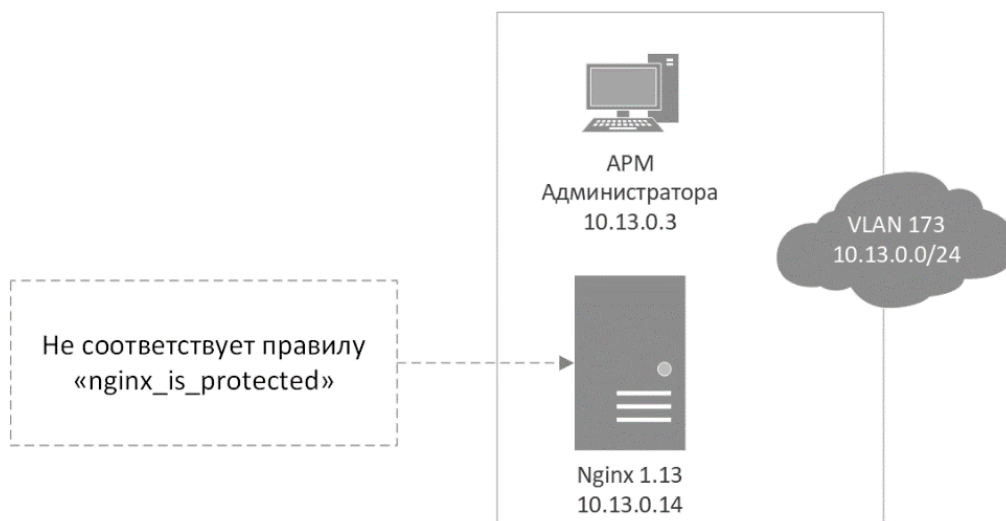


Рисунок 2. Отображение не пройденной проверки на схеме сети

Архитектура системы

В качестве архитектуры данной системы рекомендуется выбрать клиент-серверную, так как в таком случае клиентам нет необходимости:

- обновлять базу правил
- актуализировать перечень поддерживаемого ПО для построения диаграмм
- обновлять механизмы анализа схем
- устанавливать патчи для решения ошибок в работе системы
- иметь мощную комплектацию компьютера для выполнения анализа

При клиент-серверной архитектуре процедура анализа схем будет иметь следующий вид (Рисунок 3):

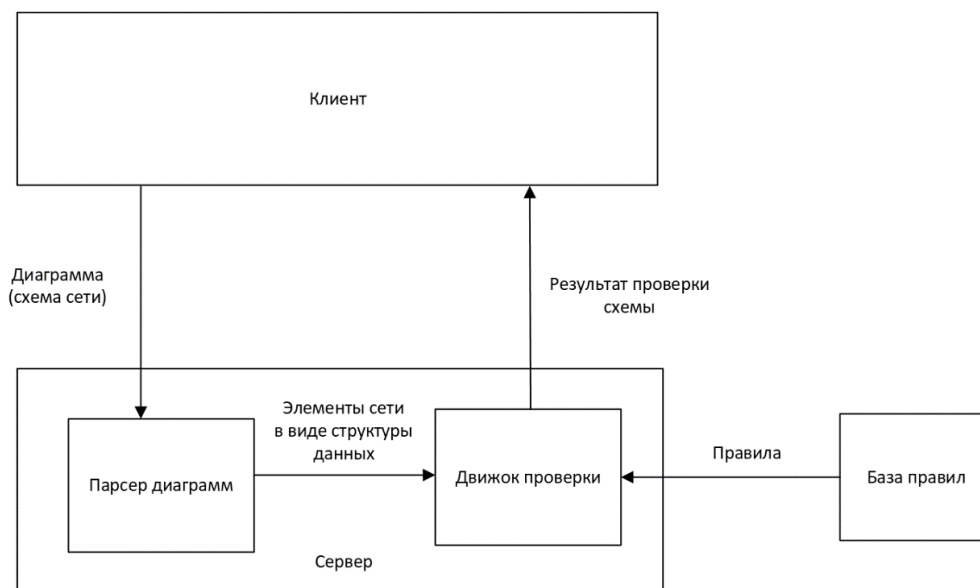


Рисунок 3. Процедура анализа схемы сети при клиент-серверной архитектуре системы

Выводы

Рассмотрена проблематика, связанная с необходимостью автоматизации процедуры проверки схем вычислительной сети. Вследствие чего было предложено разработать систему, которая предоставляет конечным пользователям возможность автоматизированной проверки схем сети на соответствие правилам, предустановленных в системе или созданных самими пользователями. Дальнейшие исследования для усовершенствования предлагаемого решения могут быть направлены на построение командной модели взаимодействия с системой, а также на использование в анализе конфигураций компонентов сети.

Литература

1. Кожевников А.Ю., Тутубалин П.И., Кирпичников А.П., Мокшин В.В. О построении подсистемы удаленного мобильного доступа к информационным ресурсам некоторой организации//Вестник Технологического университета. 2018. Т. 21. № 2. С. 139-147.
2. Кожевников А.Ю., Тутубалин П.И., Кирпичников А.П., Мокшин В.В. О разработке математических моделей, методов и программного обеспечения для проектирования перспективных изделий запрос-ответной аппаратуры//Вестник Технологического университета. 2018. Т. 21. № 2. С. 155-162.
3. Тутубалин П.И. Основные задачи прикладной теории информационной безопасности АСУ//Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2007. № 39. С. 63-72.
4. Программное управление форматами файлов Visio [Электронный ресурс]. URL:<https://docs.microsoft.com/ru-ru/office/client-developer/visio/how-to-manipulate-the-visio-file-format-programmatically> (Дата обращения 02.02.2022).

5. Сегментация сетей [Электронный ресурс]. URL:<https://www.paloaltonetworks.com/cyberpedia/what-is-network-segmentation> (Дата обращения 03.02.2022).

6. Формальные и естественные языки [Электронный ресурс]. URL:<http://csaa.ru/formalnye-i-estestvennye-jazyki/> (Дата обращения 03.02.2022).

7. Средства защиты веб-сайтов (приложений) [Электронный ресурс]. URL:<https://www.anti-malware.ru/security/application-security> (Дата обращения 14.09.2021).

СИСТЕМЫ И МЕТОДЫ ЗАЩИТЫ ИНФОРМАЦИИ

УДК 004.89

ИМПОРТОЗАМЕЩЕНИЕ И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ ОБЪЕКТОВ ТОПЛИВНО-ЭНЕРГЕТИЧЕСКОГО КОМПЛЕКСА СЕГОДНЯ И КОМПЛЕКСНАЯ БЕЗОПАСНОСТЬ В БУДУЩЕМ

IMPORT SUBSTITUTION AND INFORMATION SECURITY OF FUEL AND ENERGY COMPLEX OBJECTS TODAY AND COMPLEX SECURITY IN FUTURE

Корнеев Н.В.,
РГУ нефти и газа (НИУ) имени И.М. Губкина,
Финансовый университет при Правительстве Российской Федерации,
г. Москва, Российская Федерация

N.V. Korneev,
Gubkin Russian State University of Oil and Gas,
Financial University under the Government of the Russian Federation
Moscow, Russian Federation

e-mail: niccyper@mail.ru

Аннотация. Дан анализ ситуации с импортозамещением в секторе информационной безопасности, в связи с мировыми санкциями в отношении РФ и событиями на Украине (февраль 2022 г.). Обозначены проблемы сектора и раскрыты 3 основные причины, в контексте задач отражения киберугроз на государственные объекты критической информационной инфраструктуры РФ, и на объекты информационной инфраструктуры коммерческих организаций в условиях мировых санкциями в отношении РФ и развития событий на Украине. Выявлена системная проблема подготовки специалистов по информационной безопасности в Вузах и других научных учреждениях страны. Проведен анализ отечественных и зарубежных решений в секторе информационной безопасности, кибербезопасности и комплексной безопасности, в контексте национальной стратегии развития искусственного интеллекта на период до 2030 года. Определен набор современных отечественных решений в секторе информационной безопасности, кибербезопасности и комплексной безопасности, а также даны рекомендации для подготовки современного специалиста по информационной безопасности, на базе современных отечественных решений НПО «ЭШЕЛОН». В условиях решения задачи динамического изменения параметров безопасности критической информационной инфраструктуры компонентов экосистемы во взаимосвязи с другими компонентами раскрыта современная методология комплексной угрозы. С учетом обозначенного выше комплекса проблем сформулированы перспективные технологии обеспечения безопасности критической информационной инфраструктуры компонентов экосистемы для игроков топливно-энергетического рынка.

Abstract. An analysis of the situation with import substitution in the information security sector is given in connection with global sanctions against Russian Federation and

events in Ukraine (February 2022). The problems of the sector are identified and 3 main reasons are disclosed, in context of tasks reflecting cyber threats to state facilities of critical information infrastructure Russian Federation, and to information infrastructure facilities of commercial organizations in context of global sanctions against Russian Federation and developments in Ukraine. A systemic problem of training specialists in information security in universities and other scientific institutions of country has been identified. The analysis of domestic and foreign solutions in the sector of information security, cybersecurity and complex security was carried out in the context of the national strategy for the development of artificial intelligence for the period up to 2030. A set of modern domestic solutions in the sector of information security, cybersecurity and complex security is defined, as well as recommendations are given for the training of a modern specialist in information security, based on modern domestic solutions of NPO ESHELON. In the context of solving the problem of dynamically changing the security parameters of the critical information infrastructure of ecosystem components in conjunction with other components, a modern methodology for a complex threat is disclosed. Taking into account the above set of problems, promising technologies for ensuring the security of critical information infrastructure of ecosystem components for players in the fuel and energy market are formulated.

Ключевые слова: информационная безопасность, кибербезопасность, комплексной безопасности, искусственный интеллект, критическая информационная инфраструктура, методология комплексной угрозы.

Keywords: information security, cybersecurity, complex security, artificial intelligence, critical information infrastructure, methodology for a complex threat.

Сегодня в мире происходит значительное количество атак, и не только на государственные объекты критической информационной инфраструктуры РФ, но и на объекты информационной инфраструктуры коммерческих организаций. Вектор таких атак [1] многонаправленный, а его динамические составляющие активизировались в направлении кибератак: проникновения, взлома, DDos и т.п.

В настоящее время Российские компании не в полной мере справляются с задачами отражения таких атак, о чем свидетельствует работа специалистов на пределе своих возможностей [2]. Причин несколько.

В первую очередь это связано с тем, что объявленный еще 10 лет назад курс Правительства РФ на импортозамещение, сначала на рынке сельскохозяйственной продукции (Постановление Правительства РФ от 14.07.2012 №717 «О Государственной программе развития сельского хозяйства и регулирования рынков сельскохозяйственной продукции, сырья и продовольствия»), а затем на рынке промышленных технологий (Постановление Правительства РФ от 15.04.2014 №328 «Об утверждении государственной программы РФ «Развитие промышленности и повышение ее конкурентоспособности»), в итоге не работает. А заявления об уходе с рынка или временной приостановке работы на рынке IT-технологий РФ ряда крупных зарубежных компаний Microsoft, Adobe, DXC Technology, Oracle, Forcepoint, Fortinet, Mikrotik, Lenovo, Samsung, Dell, Spotify, HP и Cisco и др. в связи с мировыми санкциями в отношении РФ и событиями на Украине (февраль 2022 г.) лишь усиливает этот системный кризис импортозамещения. Следует сказать, что Правительство РФ в срочном порядке приняло ряд мер поддержки IT-отрасли (Указ Президента Российской Федерации от 02.03.2022 №83 «О мерах по обеспечению ускоренного развития отрасли информационных технологий в Российской Федерации») в связи с указанными событиями, но насколько они будут эффективны – покажет время.

Вторая причина связана с тем, что в РФ умеют работать, но умеют и не работать – когда заработная плата не соответствует ожиданиям, и надеются, что еще есть время сделать работу, за которую не платят. Здесь следует сказать, что российские компании занимались импортозамещением, и вели работу в этом направлении, они ее проводили: составляли планы «цифровой трансформации», моделировали «цифровые двойники», и снова составляли планы, и очень надеялись, что у них еще есть время решить вопросы импортозамещения. Мировые санкции в отношении РФ и события на Украине (февраль 2022 г.) показали, что времени уже нет. Уже достаточно давно идет информационная война, причем через дружественные нам страны [3] и отечественные решения информационной безопасности, кибербезопасности, комплексной безопасности нужны сегодня, а не через 3-5-7 лет, как написано в планах «цифровой трансформации».

Наконец, причина третья – это слабая подготовка кадров в Вузах нашей страны, и это системная проблема, начиная с 1996 года, когда была напрочь разрушена система аттестации научных кадров (кандидатов и докторов наук), изменен список рекомендуемых ВАК РФ к публикации журналов (увеличился почти 10 раз), а далее «разрушены» образовательные стандарты, в которых напрочь исчезли дидактические единицы, и так можно приводить еще много примеров. И все же основной, системной проблемой, автор здесь видит проблему отсутствия на кафедрах, в лабораториях, в руководстве Вузов и других научных организациях молодых ученых с компетенциями в области информационной безопасности, которые еще есть в РФ, и не уехали за границу [4]. Таким молодым ученым, как правило, не дают возможности развиваться в стенах указанных организаций, загружая отчетами, потом еще раз отчетами для отчетов по показателям предыдущих отчетов, корректирующими и предупреждающими действиями и т.п., после этого на научную работу времени у них не остается. Таких молодых ученых не ставят на руководящие должности, в итоге страдает вся вертикаль подготовки специалистов в области информационной безопасности, и на выходе мы получаем в лучшем случае недоученного IT-специалиста с устаревшим набором компетенций.

На чем же сегодня учить современного специалиста по информационной безопасности, и есть ли на рынке достойные отечественные решения? Ответы на эти вопросы не однозначны. Мы заведомо ограничим рассмотрения вопросов объектами ТЭК, а условия «достойные отечественные решения» будем рассматривать в контексте Национальной стратегии развития искусственного интеллекта на период до 2030 года (Указ Президента РФ от 10.10.2019 г. №490), где вопросы обеспечения безопасности критических объектов топливно-энергетического комплекса должны решаться с использованием интеллектуального анализа данных и центральную роль занимает цифровизация бизнес-процессов топливно-энергетического комплекса [5, 6].

Существующие в настоящее время решения по безопасности, в частности – информационной безопасности: IDS, DLP, log management, SIM, SEM, SIEM глубину указанной проблемы не учитывают. Попытки отдельных отечественных компаний создать такие решения: Kaspersky Industrial CyberSecurity, Kaspersky Machine Learning for Anomaly Detection от Лаборатории Касперского; КОМРАД от НПО «ЭШЕЛОН», VipNet Industrial Security от INFOTECs и др. сталкивается с отсутствием отечественных цифровых технологий в области интеллектуального анализа данных и построения систем с элементами искусственного интеллекта для целого ряда задач безопасности: физической, экономической, пожарной, информационной, психологической, безопасности интеллектуальной собственности, техногенной, безопасности от терроризма, экологической безопасности. Существующие зарубежные решения представлены на рынке более широко, например, система контроля защищенности MAXPATROL от Positive Technologies, архитектура EcoStruxure от Schneider Electric,

платформа Predix от General Electric. Они также как отечественные решения глубину указанной проблемы не учитывают.

Следует отметить в отношении платформы Predix от General Electric, в которой заявлено использование алгоритмов машинного обучения для решения задач аналитики, прежде всего – для выявления аномалий в потоках данных о состоянии промышленного оборудования. Predix следует рассматривать, как перспективный вариант решения указанной проблемы, однако здесь существует ряд объективно сложных моментов. Нет открытой информации о безопасности программных реализаций алгоритмов машинного обучения, предлагаемых Predix, что несет возможную угрозу информационной безопасности на нескольких уровнях OSI. Архитектура платформы, IoT и аналитические приложения Predix могут выполняться как в облаке, так и локально, т.е. на уровне Edge. Этот факт вызывает сомнения в отношении возможности развертывания платформы по требованиям регулятора (Приказ Министерства энергетики РФ от 06.11.2018 № 1015), который среди прочего устанавливает необходимость расположения на территории Российской Федерации центров обработки данных.

К указанным сложным моментам следует также отнести условия использования зарубежного ПО, особенно одного зарубежного вендора, что приводит к многочисленным рискам для игроков топливно-энергетического рынка и в конечном счете ставит под угрозу безопасность страны. Здесь следует выделить многочисленные юридические риски в связи с введением санкций, а также синергетического эффекта экономических потерь для России, который возникает при использовании зарубежного ПО и финансирования за счет потребителя инфраструктуры западных компаний в РФ.

В этих условиях крупным игрокам топливно-энергетического рынка приходится решать задачи цифровой трансформаций имеющейся инфраструктуры посредством создания SOK (Cybersecurity Operations Center), как отдельных рубежей защиты [7], причем как правило интегрирующих зарубежные решения, что с учетом требований импорта замещения, не всегда соответствует требованиям регуляторов. Здесь же возникает другая проблема непрерывности технологического цикла и бизнес-процессов, отказоустойчивость отдельных структурных элементов и стабильность состояния системы в целом.

В этом отношении решения от НПО «ЭШЕЛОН» являются наиболее оптимальными. Предлагаемая линейка продуктов [8] AppChecker Cloud, СКАНЕР-ВС, РУБИКОН, KOMRAD Enterprise SIEM, АК-ВС, Пульсар, ПИК ЭШЕЛОН, ГЕНЕРАТОР, СКАНЕР-ВС ИНСПЕКТОР полностью отечественного производства и позволяет закрыть вопросы информационной безопасности и кибербезопасности объектов любой инфраструктуры, в том числе критической информационной инфраструктуры предприятия.

В условиях решения задачи динамического изменения параметров безопасности критической информационной инфраструктуры компонентов экосистемы во взаимосвязи с другими компонентами следует говорить о комплексной угрозе. Комплексная угроза – угроза, состоящая из нескольких различных элементарных угроз, связанных между собой посредством определенных синхронизирующих механизмов и не обязательно существующих в одном пространстве [9].

С учетом обозначенного выше комплекса проблем перспективными технологиями обеспечения безопасности критической информационной инфраструктуры компонентов экосистемы для игроков топливно-энергетического рынка в будущем видятся следующие.

1. Формирование, цифровых двойников объектов ТЭК [10] для решения задач поддержки управления комплексной безопасностью объектов за счёт автоматизации создания компонентных моделей объектов. В основу модели объекта

может быть положен нейрограф – граф, вершины которого, представляют собой нейронные сети, моделирующие динамическое развитие объекта и окружающей его среды [11].

2. Построение предполагаемых моделей комплексных угроз. Задача может быть решена созданием системы искусственного интеллекта, обладающего знаниями о специфике формирования комплексных угроз, структуре внутренних взаимосвязей, особенностях реализации комплексных угроз и т.п. Такие знания могут быть получены только в результате обработки больших объемов данных, собранных за время работы систем мониторинга безопасности. В целом, здесь возникает круг задач, типичных для технологий Big Data, которые уже широко применяются во многих сферах, в том числе в области обеспечения защищенности и систем кибербезопасности.

3. Создание правил определения максимально схожих предполагаемых и текущих моделей комплексных угроз. Решение данной задачи включает в себя широкий спектр возможностей для применения алгоритмов интеллектуального анализа данных (Data Mining). Среди наиболее применимых в отношении данной задачи алгоритмов Data Mining можно выделить кластеризацию, классификацию и аффинитивный анализ. Не исключено использование регрессионного анализа и генетических алгоритмов.

4. Отслеживание и формирование текущей модели комплексной угрозы. По мнению автора, данная задача может быть решена созданием определенных информационно-аналитических систем, которые в том числе могут базироваться на существующих корпоративных информационных системах и средствах защиты в рамках конкретных предприятий. Интеграция и мониторинг информационных потоков, выделение критических отклонений, фиксирование событий и определение взаимосвязей методами интеллектуальной аналитики.

5. Интеллектуального обнаружения угроз в вертикально интегрированных иерархиях центров безопасности SOK, заключающееся в создании на нижних уровнях иерархии интеллектуальных систем, способных автоматизировать деятельность специалистов по информационной безопасности, взаимодействующих с действующими SIEM и IDS системами.

6. Формирование интеллектуальной системы управления комплексной безопасностью ТЭК, с учетом развития отечественной концепции системы человек-машина, как сложной информационно-энергетической системы, на базе интеллектуализации теоретических принципов физиологии и психологии человека.

7. Формирование предикативной системы, как отечественного аналога платформы Predix от General Electric с возможностью предсказывать угрозы комплексной безопасности объектов [12].

8. Разработка мультиагентов для систем мониторинга угроз функционирования и поддержки управления комплексной безопасностью объектов с интеллектуальными системами управления. В этом пункте речь идет о всех возможных физических (интеллектуальный беспилотный летательный аппарат, гейтвей, Field Agent и др) и информационных мультиагентах (сенсорах безопасности).

Выводы.

Проведен анализ отечественных и зарубежных решений в секторе информационной безопасности, кибербезопасности и комплексной безопасности, в контексте национальной стратегии развития искусственного интеллекта на период до 2030 года. Определен набор современных отечественных решений в секторе информационной безопасности, кибербезопасности и комплексной безопасности, а также даны рекомендации для подготовки современного специалиста по

информационной безопасности, на базе современных отечественных решений НПО «ЭШЕЛОН». В условиях решения задачи динамического изменения параметров безопасности критической информационной инфраструктуры компонентов экосистемы во взаимосвязи с другими компонентами раскрыта современная методология комплексной угрозы. С учетом обозначенного выше комплекса проблем сформулированы перспективные технологии обеспечения безопасности критической информационной инфраструктуры компонентов экосистемы для игроков топливно-энергетического рынка.

Литература

1. Korneev, N. (2021). The attack vector on the critical information infrastructure of the fuel and energy complex ecosystem. Paper presented at the CEUR Workshop Proceedings, 3035 59-65.
2. Копелевич И. Максим Пустовой, Positive Technologies: «Первое, с чем сейчас сталкиваются компании, – это DDoS-атаки»//БФМ.РУ: СМИ. 2022. 9 март. URL: <https://www.bfm.ru/news/494854> (дата обращения: 16.03.2022).
3. Китай обнаружил компьютерный троян, используемый разведкой США для шпионажа//ТАСС: СМИ. 2022. 14 март. URL: <https://tass.ru/mezhdunarodnaya-panorama/14068165> (дата обращения: 16.03.2022).
4. Степанова Ю. Киберзащита теряет мозг//Газета Коммерсантъ №42(7243): СМИ. 2022. 14 март. URL: <https://www.kommersant.ru/doc/5252465> (дата обращения: 16.03.2022).
5. Korneev, N. V. (2020). Intelligent complex security management system FEC for the industry 5.0. Paper presented at the IOP Conference Series: Materials Science and Engineering, 950(1) doi:10.1088/1757-899X/950/1/012016.
6. Maddikunta, P. K. R., Pham, Q. -, B, P., Deepa, N., Dev, K., Gadekallu, T. R., Liyanage, M. (2022). Industry 5.0: A survey on enabling technologies and potential applications. Journal of Industrial Information Integration, 26 doi:10.1016/j.jii.2021.100257.
7. Корнеев Н.В., Меркулов В.Д. Концепция интеллектуального обнаружения угроз в вертикально интегрированных иерархиях центров безопасности//Информационные технологии. Проблемы и решения. 2020. № 2(11). С. 134-139.
8. Разработки АО «НПО «ЭШЕЛОН»//АО «НПО «Эшелон». 2022. URL: <https://нпо-echelon.ru/production/65/> (дата обращения: 16.03.2022).
9. Korneev, N., & Merkulov, V. (2019). Intellectual analysis and basic modeling of complex threats. Paper presented at the CEUR Workshop Proceedings, 2603 23-28.
10. Korneev, N., & Merkulov, V. (2022). Security concept of digital twins of industrial fuel and energy complex objects doi:10.1007/978-981-16-3844-2_33.
11. Korneev, N.V. (2019). A neurograph as a model to support control over the comprehensive objects safety for BIM technologies. Paper presented at the IOP Conference Series: Earth and Environmental Science, 224(1) doi:10.1088/1755-1315/224/1/012021.
12. Korneev, N.V., Korneeva, J.V., Yurkevichyus, S.P., & Bakhturin, G.I. (2022). An approach to risk assessment and threat prediction for complex object security based on a predicative self-configuring neural system. Symmetry, 14(1) doi:10.3390/sym14010102