ISSN 2225-5435

Вестник УрФО.

БЕЗОПАСНОСТЬ В ИНФОРМАЦИОННОЙ СФЕРЕ

№ 2(52) / 2024

Подписной индекс 73852 в каталоге «Почта России»

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство ПИ № ФС77-65765 от 20.05.2016

Адрес редакции и издателя: Россия, 454080, г. Челябинск, пр. Ленина, д. 76. ЮУрГУ, Издательский центр
Тел./факс (351) 267-97-01.

16+

Электронная версия журнала в Интернете:
www.info-secur.ru,
e-mail: urvest@mail.ru

Journal of the Ural Federal District.

ISSN 2225-5435

INFORMATION SECURITY

№ 2(52) / 2024

Subscription index 73852 in the «Russian Post» catalog

The journal is registered by the Federal service in the field of communication, information technology and mass communications.

Certificate
PI No. ФС77-65765 dd. 05/20/2016

Editorial and publisher address: Russia, 454080, Chelyabinsk, Lenin Avenue, 76 SUSU, Publishing Center
Phone / fax (351) 267-97-01.

16+

Electronic version of the magazine in the Internet:
www.info-secur.ru,
e-mail: urvest@mail.ru

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

IN THIS ISSUE

SYSTEM ANALYSIS, MANAGEMENT AND INFORMATION PROCESSING

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

УДК 004.042, 004.89 Вестник УрФО № 2(52) / 2024, с. 5–12

Крыжановская Ю. А., Курченкова Т. В.

10.14529/secur240201

ПРИМЕНЕНИЕ ДЕРЕВЬЕВ КЛАССИФИКАЦИИ ДЛЯ РЕАЛИЗАЦИИ СКОРИНГОВОЙ СИСТЕМЫ

В статье обсуждаются результаты разработки системы оценивания кредитоспособности заявителя. На основе анализа существующих моделей выбран подход применения деревьев классификации, преимуществам которого являются минимальное количество ручной конфигурации, возможность обработки данных с числовыми и категориальными атрибутами, а также наглядная интерпретация причин принятия того или иного решения. Модель классификации обучалась на общедоступном наборе данных, предоставляющем данные по 20 параметрам. Для хранения данных приложения была спроектирована база данных. Разработанное приложение основано на клиент-серверной архитектуре и реализует паттерн MVC.

Ключевые слова: *скоринговая система, деревья классификации, база данных, модель классификации, принятие решений, клиент-серверное приложение.*

Kryzhanovskaya Yu. A., Kurchenkova T. V.

CLASSIFICATION TREESBASED SCORING SYSTEM DEVELOPMENT

The article discusses the developed system for assessing the borrower's creditworthiness. Based on the existing models analysis, the approach of using classification trees was chosen. It's advantages are the minimum manual configuration amount, the ability to process data with numerical and categorical attributes, as well as visual interpretation reasons for making a particular decision. The classification model was trained on a publicly available dataset providing data on 20 parameters. A database was designed to store application data. The developed application is based on client-server architecture and implements the MVC pattern.

Keywords: *scoring system, classification trees, database, classification model, decision making, client-server application.*

Введение

На протяжении истории существования банковской системы недобросовестные заемщики были большой проблемой, приносящей значительные убытки. И даже опытные кредитные эксперты [1] далеко не всегда могут с достаточной точностью оценить потенциального заемщика. Для решения этой проблемы применяются системы, способные на основании информации о предыдущих кредитных историях делать гораздо более точные прогнозы [2], тем самым превентивно решая потенциальные проблемы с недобросовестным клиентом. Также использование автоматизированных систем позволяет не предъявлять высоких требований к квалификации кредитного инспектора и влияет на скорость рассмотрения заявок.

Целью данной работы является разработка клиент-серверного приложения кредитного скоринга на основе деревьев классификации, дающего оценку заемщика по данным из его анкеты. Основное назначение приложения – быстрая оценка заемщика и сокращение времени на изучение документов.

Приложение должно обладать такими функциональными возможностями, как возможность построения модели принятия решений на основе подготовленного датасета, ввод данных о клиенте в анкету, сохранение переданных данных в базе данных, вынесение решения о предоставлении кредита на основе переданных в анкете данных, автоматическое перестроение модели через определенный системным администратором промежуток времени на основе сохраненных в БД анкет.

Анализ существующих моделей классификации

Для создания системы кредитного скоринга необходимого выбрать математическую модель её реализации. При выборе модели для скоринговых систем оценки кредитного риска заемщиков важно учитывать различные факторы, такие, как размер и качество данных, сложность данных, требуемая скорость и точность предсказаний, а также стоимость и время разработки. На первом этапе работы были проанализированы различные подходы к данной проблеме [3, 4].

Один из популярных и хорошо изученных подходов – метод ближайших соседей [5]. По сравнению с другими подходами характеризуется относительно простой реализацией и изученностью. Однако, если число признаков, на которых строится классификация, велико, метод плохо работает, а для признаков может быть сложным подобрать веса и определить какие из них являются важными.

Еще один применяемый для построения скоринговых систем подходов – логистическая регрессия [6]. Это модель хорошо подходит для бинарных классификаций, а с многоклассовыми задачами менее эффективна. Также важна подготовка данных, причем для стабильности получаемого результата может потребоваться большой объем выборки.

В скоринговых системах применим и нечетко-множественный подход [7]. К преимуществам нечеткой модели можно отнести способность оперировать одновременно количественными и качественными характеристиками, к недостаткам – сложность интерпретации результатов и трудности в обновлении и адаптации, а также худшую производительность.

Также в скоринговых системах возможно применение генетических алгоритмов [8]. В этом случае генерируется начальное множество скоринговых функций, к которому применяются операции скрещивания и мутации с выводом из рассмотрения наименее пригодных функций. Генетические алгоритмы применимы в системах оценки кредитоспособности заемщиков благодаря гибкости и надежности, однако требуют значительных вычислительных ресурсов и времени для нахождения оптимальных решений, а также тщательной настройки параметров.

Еще один метод построения скоринговой модели – нейронная сеть [9]. Эта модель может обрабатывать сложные данные и выявлять нелинейные зависимости, но требует больших вычислительных ресурсов и времени обучения.

Также возможно применение модели, основанной на использовании деревьев классификации [10]. Такая модель позволяет отнести объект – потенциального заемщика – к одному из заранее известных классов, например, к классу «Одобрить заявку» или к классу «Отклонить заявку». Это позволяет, основываясь на значениях признаков, разделить данные на подмножества и найти лучшие правила классификации. В результате получается представление правил в виде иерархической структуры, где каждому объекту соответствует единственный узел-решение. Модель допускает адаптацию к меняющимся условиям путем перестроения дерева.

Алгоритм классификации

Для реализации скоринговой системы, представленной в данной работе, из перечисленных моделей был выбран алгоритм дерева решений, так как, как сказано выше, дерево может перестраиваться при добавлении новых данных и игнорировать несущественные признаки. К преимуществам алгоритма можно отнести и минимальное количество ручной. Также он удобен подходит для случает обработки данных со смешанными атрибутами (числовыми и категориальными). Еще одним достоинством является наглядная интерпретация причин принятия того или иного решения с помощью дерева.

Одной из ключевых составляющих при построении дерева решения является энтропия Шеннона (информационная энтропия) — мера неопределённости, связанной со случайной величиной. Эквивалентно, энтропия Шеннона – мера информационной упорядоченность, однородности множества.

Построение дерева начинается с задания множества объектов, предикатов и минимального числа объектов для поиска разбиения. Каждый объект имеет набор атрибутов со значениями. Предикаты являются логическими условиями (больше, меньше, равно и т. д.). При построении рекурсивно вычисляется наилучшее разбиение исходного множества. Для каждого объекта множества ищется такое сочетание атрибута и предиката, которое приведет к уменьшению среднего значения энтропии Шеннона. Критерием остановки поиска разбиения является либо близкое к нулю значение энтропии, либо размер оставшегося после разбиения множества, меньший заданного минимального. Из полученного множества путем подсчета наиболее часто встречающегося класса объекта создается лист дерева с полученным классом.

После завершения построения производится оптимизация полученного дерева. Рекурсивно вверх, начиная с нижних узлов дерева, если оба листа относятся к одному классу, то они сливаются в 1 лист на место родительского узла. Чтобы определить к какому классу относится объект при помощи дерева классификации, следует рекурсивно спускаться по дереву. Направление при этом выбирается исходя из значений предикатов, применяемых к классифицируемому объекту. С точки зрения интерпретации результатов, каждый путь от корня дерева до листа отражает объяснение того, почему объект был отнесён к указанному классу. Проблема переобучения может быть решена при помощи леса решений, который включает несколько деревьев. При этом результат классификации определяется «голосованием», при котором ответом выбирается тот класс, за которой проголосовало большее число деревьев.

В представленной работе для прогнозирования строится лес из нечетного количества деревьев. При классификации обрабатываемого объекта каждое дерево выносит свое решение о том, к какому классу будет отнесен объект. Определить, с какой вероятностью объект принадлежит конкретному классу, можно исходя из количества проголосовавших за этот результат деревьев. При этом каждое дерево обучается на своём случайном подмножестве исходной обучающей выборки данных, что необходимо для уменьшения числа ошибочных классификаций. Слишком большое количество деревьев в лесе не имеет практической целесообразности, так как увеличение количества деревьев больше, чем до 11, практически не увеличивает качество классификации.

Средства реализации

В качества языка программирования в работе используется Java 17.0.3. Выбор обусловлен большим количеством фреймворков и библиотек для разработки web-приложений, а также хорошей скоростью работы. В частности, Spring Framework предлагает функцию внедрения зависимостей, которая позволяет объектам определять свои собственные зависимости, которые контейнер Spring позже вводит в них. Также в работе используется Spring Boot, упрощающий разработку web-приложений благодаря возможности автоматической конфигурации с помощью предустановленных зависимостей, которые не нужно настраивать вручную. Spring Data — дополнительный удобный механизм для взаимодействия с сущностями базы данных, организации их в репозитории, извлечения и изменения данных.

В качестве системы управления базами данных (СУБД) была выбрана PostgreSQL. СУБД отличается высокой надёжностью и хорошей производительностью, поддерживает транзакции (ACID), поддерживает хранение больших двоичных объектов, а репликация реализована встроенными механизмами. Также PostgreSQL часто используется в связке

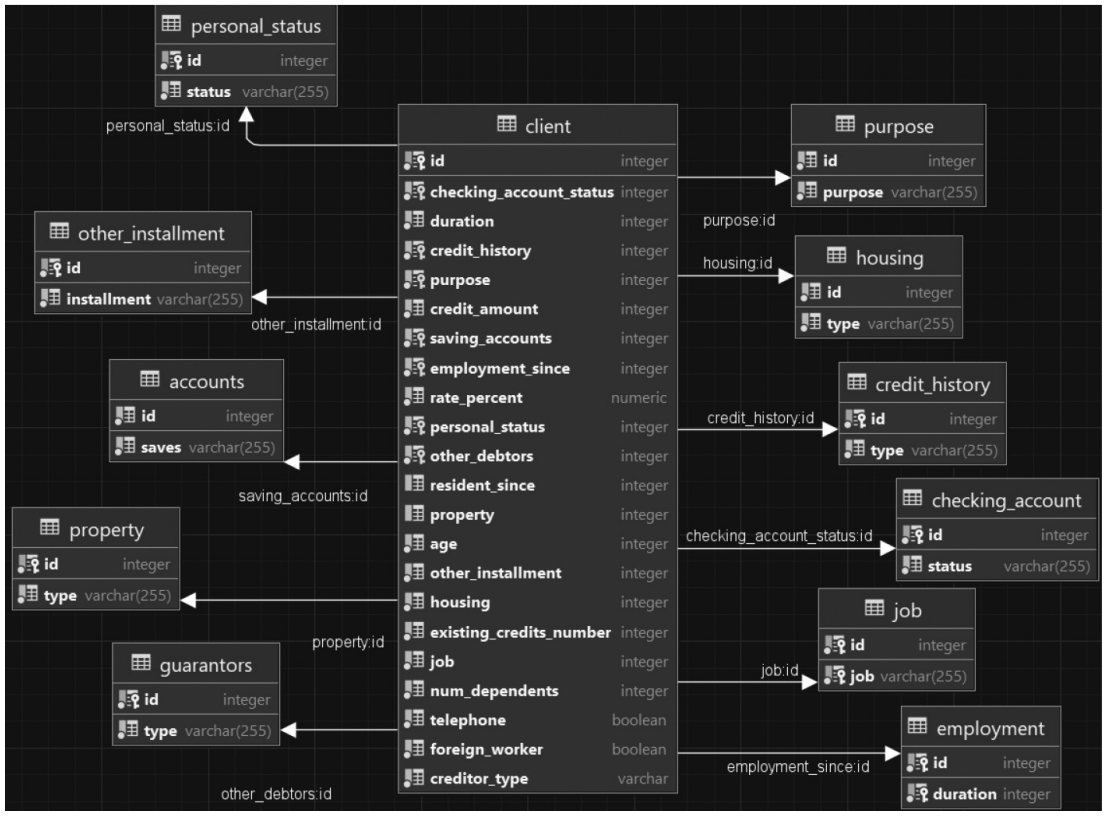

Рис. 1. Схема БД

с языком программирования Java, благодаря чему существует большая база знаний по совместной работе этих продуктов.

Структура базы данных

Для скоринговой системы была разработана база данных (БД), которая включает 12 таблиц:

- client содержит основную анкетную информацию о клиенте, которая используется для переобучения модели и оценки возможности предоставления кредита;
- checking_account хранит статусы расчетного счета;
- purpose – существующие цели кредита;
- credit_history – хранение статусов кредитной истории клиента;
- accounts хранит в себе статусы сберегательного счета;
- personal_status – социальные статусы и пол клиента;
- other_installment – статусы расчетного счета;
- property отражает существующие типы собственности;
- guarantors содержит возможные типы прочих должников и поручителей;
- housing хранит существующие виды владения жильем;
- job – типы рабочей квалификации клиента;
- employment – существующие типы стажа работы.

Структура БД показана на рис. 1.

Обучение модели

Одним из важнейших этапов создания модели классификации, определяющим качество последующих предсказаний, является подбор качественной обучающей выборки данных. Подобранный датасет должен содержать достаточное количество данных, быть применимым в реальной системе и учитывать ключевые характеристики заемщика.

Для обучения был подобран общедоступный датасет из репозитория машинного обучения UCI (University of California, Irvine). Датасет содержит 5000 обезличенных реальных кредитных историй за 2008 год и классификацию того, считается ли заявитель хорошим или плохим. Он предоставляет данные по 20 параметрам как числовым, так и категориальным, которые нашли отражение в структуре таблицы client.

Структура приложения

Разрабатываемое приложение основано на клиент-серверной архитектуре и реализует паттерн MVC. На стороне клиента происходит получение запрошенных данных, а также ввод данных пользователем в систему. Сервер получает запросы с клиентской части и реагирует на них. Осуществляются запросы к базе данных с последующей необходимой обработкой над полученной выборкой данных. Впоследствии эти данных заполняются в представление и передаются в виде ответа на клиентскую часть. База данных позволяет сохранять данные в ПЗУ, тем самым предотвращая их потерю и не засоряя оперативную память сервера.

Работа приложения начинается с http-запроса пользователя. Spring перенаправляет выполнение поступающих http-запросов на соответствующие методы контроллеров. Классы контроллеров содержат реализации service-классов. В сервисном классе производится обращение к деревьям решений для классификации переданной анкеты, а также запись анкеты в базу данных.

На начальном этапе разработки приложения были определены основные доменные сущности и классы. Модуль классификации состоит из классов, представленных на рис. 2.

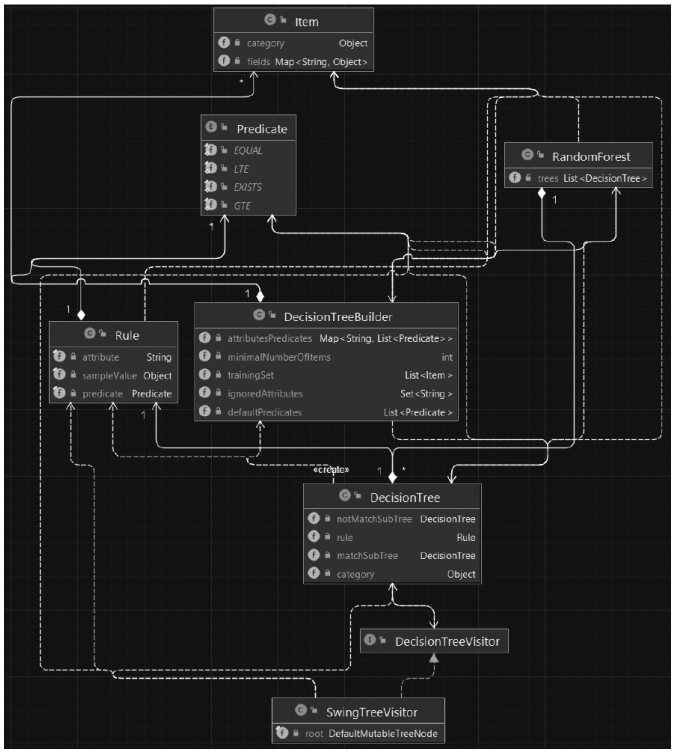

Рис. 2. UML диаграмма классов модуля классификации

Класс Item предназначен для хранения значения атрибутов классифицируемого объекта. Для большей гибкости он содержит в себе ассоциативный массив, ключом в котором является название атрибута, а значением – значение атрибута. Это позволяет легко добавлять и убирать поля при необходимости, не затрагивая остальные компоненты системы.

Класс Predicate хранит в себе логические выражения, представленные в виде полей класса. Каждое поле содержит в себе переопределенный метод «Predicate», принимающий 2 сравниваемых значения и возвращающий логический результат его вычисления. Таким образом, достигается инкапсуляция работы с логическими выражениями в отдельных объектах, благодаря чему можно легко настраивать необходимый набор предикатов для каждого атрибута объекта.

Класс Rule хранит сочетание атрибута объекта, его значения и предиката. Именно это сочетание определяет искомое в процессе построения дерева разбиение множества на подмножества, удовлетворяющее и не удовлетворяющее этому правилу. Впоследствии оно сохраняется в узле дерева, и классифицируемый объект проходит проверку на соответствие этому правилу.

Класс DecisionTree непосредственно отвечает за хранение и построения дерева классификации. Для этого в нем определены поля того же типа matchSubTree и notMatchSubTree, что делает структуру данных рекурсивной. Поле rule предназначено для хранения правила типа «Rule» и инициализировано, только если текущий объект является узлом дерева. Поле category отображает класс объекта и заполнено только когда текущий объект является листом дерева.

Класс DecisionTreeBuilder является вспомогательным для построения дерева. Он хранит в себе параметры, которые нужно только для построения и не должны сохраняться в самом дереве.

Класс RandomForest предназначен для хранения леса классификации. Он хранит в себе список деревьев и методы для разбиения

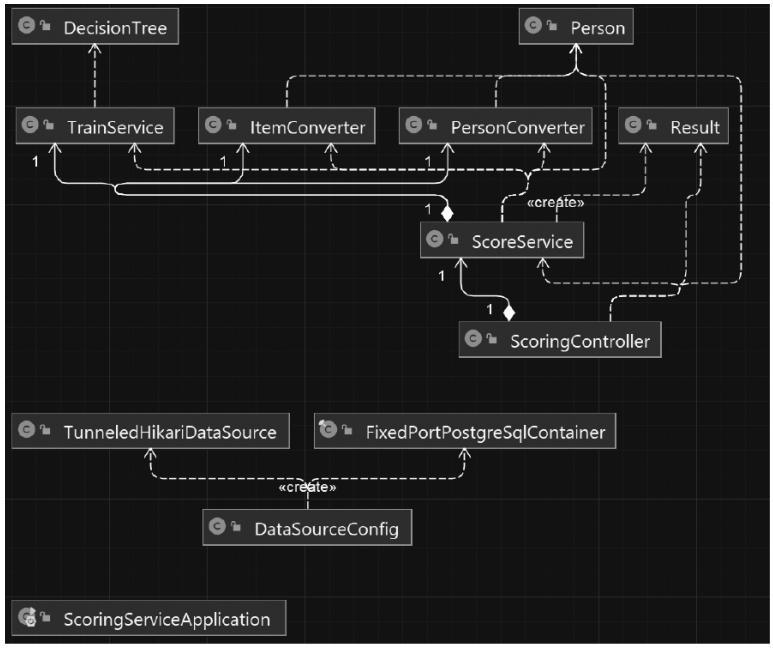

Рис. 3. UML диаграмма классов модуля web части приложения

исходного датасета на случайные обучающие подмножества.

Класс SwingTreeVisitor предназначен для обхода дерева с последующим его выводом на экран.

Общая структура приложения представлена на рис. 3.

На диаграмме видна структура web части приложения. Класс ScoringController принимает приходящие запросы. Классы ScoreService и TrainService являются сервисными и отвечают за классификацию поступающей анкеты и переобучение модели соответственно. Классы Person и Result отображают модели клиента и результата классификации.

Также реализованы классы конвертеров, для перевода поступающих данных в формат, обрабатываемый деревьями классификации. Также в конфигурационных классах заданы настройки подключения к базе данных и создание докер-образа приложения.

Общее описание приложения

Перед стартом приложения в конфигурационном файле application.properties необходимо указать путь к файлу с обучающим датасетом, количество генерируемых деревьев в лесе, а также настройки для подключения к базе данных. Далее при старте генерируется деревья классификации на основе переданного датасета. При желании, в конфигурационном файле можно включить отображение всех построенных деревьев, которые различаются между собой, поскольку обучаются на случайных подмножествах исходного датасета.

Для ввода данных предусмотрена анкета, находящаяся на главной странице приложения. При нажатии на кнопку «Получить оценку» на сервер отправляется HTTP POST запрос с переданными данными на адрес /score. На сервере запрос принимает контроллер, где он валидируется. Далее, в сервисном слое, вызывается метод, обрабатывающий данные клиента через деревья классификации и возвращающий количество решений «За» и «Против», на основе которых формируется окончательный результат. После этого анкета с принятым решением сохраняется в базе данных. Далее результат возвращается контроллером и отображается на странице ввода.

Для переобучения предусмотрен HTTP PATCH метод /retrain, с помощью которого можно переобучить модель, дополнив исходный датасет накопленными данными из БД.

Заключение

В результате проделанной была работы спроектирована база данных и реализовано приложение, позволяющее с помощью деревьев классификации выносить решение о возможности предоставления заемщику кредита, исходя из данных анкеты. Функциональность приложения может быть расширена, также есть возможность масштабирования. Приложение было протестировано с использованием 200 записей датасета, не применявшихся при обучении модели, и показало результат классификации, совпадающий с находящимся в датасете на 85%, что позволяет сделать вывод о применимости реализованного подхода в реальной системе.

Литература

1. Hutchins J. The US Farm credit system and agricultural development: Evidence from an early expansion, 1920–1940. American Journal of Agricultural Economics. 2023. Vol. 105. No. 1. P. 3-26. DOI: 10.1111/ajae.12290/.

2. Pushkareva L.V., Galochkina O.A., Bezgacheva O.L. Current trends in the banking system of Russia. Espacios. 2019. Vol. 40. No. 4. P. 22-29.

3. Addo, P., Guegan, D., & Hassani, B. (2018). Credit Risk Analysis Using Machine and Deep Learning Models. Risks, 6(2), 38. doi: 10.3390/risks6020038

4. Munkhdalai, L., Munkhdalai, T., Namsrai, O., Lee, J., & Ryu, K. (2019). An Empirical Comparison of Machine-Learning Methods on Bank Client Credit Assessments. Sustainability, 11(3), 699. doi: 10.3390/su11030699

5. Алексеева В. А. Применение метода ближайших соседей при моделировании кредитных рисков / В. А. Алексеева, Р. И. Калимулинна // Вестник Ульяновского государственного технического университета. – 2014. - № 3. – С. 54-56

6. Сорокин А. С. Построение скоринговых карт с использованием модели логистической регрессии / А. С. Сорокин // НАУКОВЕДЕНИЕ. – 2014. - № 2.

7. Волкова Е.С., Гисин В.Б., Соловьев В.И. Методы теории нечетких множеств в кредитном скоринге // Финансы и кредит. – 2017. – Т. 23, № 35. – С. 2088 – 2106. https://doi.org/10.24891/fc.23.35.2088

8. Ong C. S., Huang J. J., Tzeng G. H. Building credit scoring models using genetic programming // Expert Systems with Applications. – 2005. – Vol. 29. – [1] 1. – P. 41-47.

9. Кадиев А.Д., Чибисова А.В. Нейросетевые методы решения задачи кредитного скоринга. Математическое моделирование и численные методы, 2022, № 4, с. 81–92.

10. Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research //European Journal of Operational Research. – 2015. – Vol. 247. – [1] 1. – P. 124-136.

References

1. Hutchins J. The US Farm credit system and agricultural development: Evidence from an early expansion, 1920–1940. American Journal of Agricultural Economics. 2023. Vol. 105. No. 1. P. 3-26. DOI: 10.1111/ajae.12290/.

2. Pushkareva L.V., Galochkina O.A., Bezgacheva O.L. Current trends in the banking system of Russia. Espacios. 2019. Vol. 40. No. 4. P. 22-29.

3. Addo, P., Guegan, D., & Hassani, B. (2018). Credit Risk Analysis Using Machine and Deep Learning Models. Risks, 6(2), 38. doi: 10.3390/risks6020038

4. Munkhdalai, L., Munkhdalai, T., Namsrai, O., Lee, J., & Ryu, K. (2019). An Empirical Comparison of Machine-Learning Methods on Bank Client Credit Assessments. Sustainability, 11(3), 699. doi: 10.3390/su11030699

5. Alekseyeva V. A. Primeneniye metoda blizhayshikh sosedey pri modelirovanii kreditnykh riskov / V. A. Alekseyeva, R. I. Kalimulinna // Vestnik Ul'yanovskogo gosudarstvennogo tekhnicheskogo universiteta. – 2014. - № 3. – S. 54-56

6. Sorokin A. S. Postroyeniye skoringovykh kart s ispol'zovaniyem modeli logisticheskoy regressii / A. S. Sorokin // NAUKOVEDENIYe. – 2014. - № 2.

7. Volkova Ye.S., Gisin V.B., Solov'yev V.I. Metody teorii nechetkikh mnozhestv v kreditnom skoringe // Finansy i kredit. – 2017. – T. 23, № 35. – S. 2088 – 2106. https://doi.org/10.24891/fc.23.35.2088

8. Ong C. S., Huang J. J., Tzeng G. H. Building credit scoring models using genetic programming // Expert Systems with Applications. – 2005. – Vol. 29. – [1] 1. – P. 41-47.

9. Kadiyev A.D., Chibisova A.V. Neyrosetevyye metody resheniya zadachi kreditnogo skoringa. Matematicheskoye modelirovaniye i chislennyye metody, 2022, № 4, s. 81–92.

10. Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research //European Journal of Operational Research. – 2015. – Vol. 247. – [1] 1. – P. 124-136.

КРЫЖАНОВСКАЯ Юлиана Александровна, старший преподаватель, кафедра Кибербезопасности информационных систем, факультет Прикладной математики, информатики и механики, федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». 394018, г. Воронеж, Университетская площадь, 1. E-mail: jak@mail.ru

КУРЧЕНКОВА Татьяна Викторовна, кандидат технических наук, доцент, доцент кафедры Программного обеспечения и администрирования информационных систем, факультет Прикладной математики, информатики и механики, федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». 394018, г. Воронеж, Университетская площадь, 1. E-mail: tatyana36136@mail.ru

KRYZHANOVSKAYA Yuliana Alexandrovna, Senior Lecturer, Department of Cybersecurity of Information Systems, Faculty of Applied Mathematics, Computer Science and Mechanics, Federal State Budgetary Educational Institution of Higher Education «Voronezh State University». 394018, Voronezh, University Square, 1. E-mail: jak@mail.ru

KURCHENKOVA Tatyana Viktorovna, Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Software and Administration of Information Systems, Faculty of Applied Mathematics, Computer Science and Mechanics, Federal State Budgetary Educational Institution of Higher Education «Voronezh State University». 394018, Voronezh, University Square, 1. E-mail: tatyana36136@mail.ru

УДК 004.93 Вестник УрФО № 2(52) / 2024, с. 13–23

Котенко В. В., Румянцев К. Е., Хаджиева Л. К. DOI: 10.14529/secur240202

ОЦЕНКА ВОЗМОЖНОСТЕЙ НЕЙРОЛИНГВИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Приведены результаты исследований возможностей нейролингвистической текстовой идентификации систем искусственного интеллекта (СИИ). Для достижения целей исследования применён апробированный программный комплекс оценки информационных параметров нейролингвистической текстовой идентификации интеллектуальных систем. Результаты исследований указывают на возможность применения в качестве параметров нейролингвистической идентификации информационных характеристик текстовой составляющей систем искусственного интеллекта при рассмотрении их в качестве источников информации. Приведены зависимости основных и производных параметров нейролингвистической текстовой идентификации от изменений систем искусственного интеллекта. Проанализированы изменения показателей систем искусственного интеллекта по таким основным (информационная емкость, энтропия, избыточность) и производным (коэффициент избыточности, коэффициент вербальности) параметрам. По результатам анализа установлено, во-первых, что значения основных и производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта при постановке одной и той же задачи разным СИИ значительно отличаются. Во-вторых, информационные характеристики текстовой составляющей систем искусственного интеллекта можно использовать в качестве параметров нейролингвистической текстовой идентификации систем искусственного интеллекта. В-третьих, в качестве преимущественных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта целесообразно использовать из основных параметров информационную емкость и избыточность, из производных параметров коэффициент вербальности, поскольку наблюдается максимальное изменение показателей данных параметров при переходе от одной СИИ к другой (в среднем на 40 %). Полученные результаты открывают возможность новых подходов к обнаружению идентификации и аутентификации систем искусственного интеллекта при решении проблем противодействия информационным вторжениям, несущем угрозы функционированию компьютерных сетей и телекоммуникационным систем.

Ключевые слова: *искусственный интеллект, система искусственного интеллекта, нейролингвистическая идентификация, избыточность, вербальность, информационная емкость.*

NEURO-LINGUISTIC IDENTIFICATION OF ARTIFICIAL INTELLIGENCE SYSTEMS

The results of research on the possibilities of neuro-linguistic textual identification of artificial intelligence (AI) systems are presented. To achieve the objectives of the study, an approved software package for evaluating the information parameters of neuro-linguistic textual identification of intelligent systems was used. The research results indicate the possibility of using the information characteristics of the textual component of artificial intelligence systems as parameters of neuro-linguistic identification when considering them as sources of information. The dependences of the basic and derived parameters of neuro-linguistic textual identification on changes in artificial intelligence systems are presented. The changes in the indicators of artificial intelligence systems according to such basic (information capacity, entropy, redundancy) and derivative (redundancy coefficient, verbality coefficient) parameters are analyzed. According to the results of the analysis, it was found, firstly, that the values of the basic and derived parameters of the neuro-linguistic textual identification of artificial intelligence systems when setting the same task differ significantly from one another. Secondly, the information characteristics of the textual component of artificial intelligence systems can be used as parameters of the neuro-linguistic textual identification of artificial intelligence systems. Thirdly, as the predominant parameters of neuro-linguistic textual identification of artificial intelligence systems, it is advisable to use information capacity and redundancy from the main parameters, and the verbality coefficient from the derived parameters, since there is a maximum change in the indicators of these parameters during the transition from one SII to another (on average by 40%). The results obtained open up the possibility of new approaches to the detection of identification and authentication of artificial intelligence systems in solving problems of countering information intrusions that pose threats to the functioning of computer networks and telecommunications systems.

Keywords: *artificial intelligence, artificial intelligence system, neuro-linguistic identification, redundancy, verbality, information capacity.*

Введение

Стремительное развитие искусственного интеллекта (ИИ) и его использование практически во всех областях жизнедеятельности человека нуждается в изучении и исследовании. С одной стороны развитие искусственного интеллекта прокладывает путь к новой эре технологических достижений, которые существенно повлияют на нашу повседневную жизнь и улучшат наш быт, с другой стороны такое развитие несет большую опасность. Использование ИИ злоумышленниками в своих целях. ИИ может хранить или обрабатывать информацию, к которой может получить доступ злоумышленник.

Отсюда следует важность решения проблем идентификации и обнаружения систем искусственного интеллекта (СИИ).

Нейролингвистика — отрасль психологической науки, пограничная для психологии, неврологии и лингвистики, изучающая «мозговые механизмы речевой деятельности [21].

Нейролингвистическая идентификация — это идентификация интеллектуальных систем по параметрам идентификатора, отражающего мозговые механизмы речевой деятельности.

Результатом речевой деятельности может выступать текст, что объясняет существование нейролингвистической текстовой идентификации.

Установлено, что фактором идентификации может являться язык речевой деятельности.

Исследования, проведённые авторами, показали, что проблема идентификации мозговых механизмов речевой деятельности может быть решена при использовании в качестве идентификатора количества информации, содержащегося в логических элементах речи (словах, слогах и т.п.).

В качестве параметров идентификатора в данном случае выступают: коэффициент избыточности, коэффициент вербальности, эмпирическая энтропия, информационная емкость.

Проведённые авторами исследования показали результативность данного подхода при идентификации интеллектуальных систем (людей). При исследовании использовались тексты по одинаковой тематике, произвольно формируемой различными интеллектуальными системами. Применение предложенной нейролингвистической текстовой идентификации показало возможность довольно точной идентификации интеллектуальных систем.

В статье решалась проблема применения разработанного подхода к идентификации систем искусственного интеллекта. Под системой искусственного интеллекта понималась общепринятое определение: «Системы искусственного интеллекта – это технологии, позволяющие компьютерам и программам выполнять задачи, которые раньше были доступны только человеку» [22]. В данном случае рассматривалась задача дублирование мозговой деятельности в части формирования речевого текста по определенной тематике. Исходя из этого в качестве систем искусственного интеллекта рассматривались системы формирования текста по заданной тематике.

Постановка задачи

Ставилась задача оценки возможности применения разработанного подхода для нейролингвистической текстовой идентификации систем искусственного интеллекта. Успешное решение этой задачи позволяло показать индивидуальность систем искусственного интеллекта откуда следовало существование задатков моделирования мыслительной деятельности.

Целью исследования является анализ нейролингвистической текстовой идентификации систем искусственного интеллекта и определение вероятности использования параметров в качестве факторов нейролингвистической идентификации систем искусственного интеллекта.

Задачами исследования являются:

1. Количественная оценка отличий значений основных и производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта при постановке одной и той же задачи разным системам ИИ.

2. Анализ основных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта.

3. Анализ производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта.

4. Определение средних значений основных и производных параметров нейролингвистических текстовых идентификаторов систем искусственного интеллекта.

Описание исследований

В исследовании проанализированы следующие системы искусственного интеллекта: Gerwin, YandexGPT, RoboGPT, CopyMonkey, Micrasoft Copilot, Всезнайка, ChatGPT 3.5, для которых приняты сокращения СИИ-1, СИИ-2, …, СИИ-n. Всем исследуемым системам ИИ определена одна и та же задача для решения. Постановка задачи формулируется на русском языке. Первоначально предполагается, что при постановке одной и той же задачи на одном языке разные системы ИИ должны выдавать одинаковый результат. Однако, исследование показывает, что результат каждой системы ИИ индивидуален и отличается от всех остальных [2, 3]. Полученные результаты всех исследуемых систем ИИ преобразуются в нужный формат и анализируются с помощью апробированного программного комплекса оценки информационных параметров нейролингвистической текстовой идентификации интеллектуальных систем. Показатели всех параметров меняются от одной системы ИИ к другой [4 - 13]. Интерфейс апробированного программного комплекса оценки параметров нейролингвистической текстовой идентификации систем искусственного интеллекта приведен на рис. 3.

Программный комплекс при обработке данных разных СИИ выдает результаты основных показателей: информационной емкости, энтропии и избыточности.

Информационная емкость ансамбля дискретного источника определяется из числа элементов выборочного пространства K по формуле:

$$H_{max}[U] = log_2 K \quad (1)$$

Эмпирическая энтропия определяется в виде:

$$H_L(U) = M\left[J^*[u_i]\right] \quad (2)$$

где $J^*[u_i]$ оценка количества информации в сообщении u_i.

В общем виде избыточность дискретных источников информации определяется как:

$$B[U] = H_{max}[U] - H[U] \quad (3)$$

где $H_{max}[U]$ – максимально возможная энтропия выборочного пространства ансамбля U источника; $H[U]$ – энтропия источника.

Для расчёта производных параметров коэффициента избыточности μ_B и коэффициента вербальности G_B использованы формулы:

$$\mu_B = \frac{H_{Bmax} - H_B}{H_{Bmax}} \quad (4)$$

$$G_B = \frac{H_B}{H_{Bmax} - H_B} \quad (5)$$

где H_{Bmax} - среднее значение информационной емкости;

H_B - среднее значение энтропии.

Коэффициент избыточности показывает, какая доля возможной информационной емкости на букву алфавита не используется алфавитом. Так как является безразмерной величиной, то его обычно измеряют в процентах [14 - 17].

Сводка дискретных алгоритмов фильтрации:

Модель сообщения:

$$\boldsymbol{J_x}(j+1) = \boldsymbol{\Phi}(j+1, j)\boldsymbol{J}_x(j) + \boldsymbol{\Gamma}(j)\boldsymbol{J}_w(j) \quad (6)$$

Модель наблюдения:

$$\boldsymbol{J}_z(j) = \boldsymbol{H}(j)\boldsymbol{J_x}(j) + \boldsymbol{J}_v(j) \quad (7)$$

Априорные данные:

$$\mathrm{E}\{\boldsymbol{J_w}(\boldsymbol{j})\} = 0; \mathrm{E}\{\boldsymbol{J_v}(j)\} = 0; \mathrm{E}\{\boldsymbol{J_x}(0)\} = \mu_{\boldsymbol{x}}(0)$$

$$cov\{\boldsymbol{J_w}(j), \boldsymbol{J_w}(k)\} = \boldsymbol{V}_{\boldsymbol{J_w}}(j)\delta_K(j-k)$$

$$cov\{\boldsymbol{J_v}(j), \boldsymbol{J_v}(k)\} = \boldsymbol{V}_{\boldsymbol{J_v}}(j)\delta_K(j-k)$$

$$cov\{\boldsymbol{J_w}(j), \boldsymbol{J_v}(k)\} = cov\{\boldsymbol{J_x}(0), \boldsymbol{J_v}(k)\} =$$

$$= cov\{\boldsymbol{J_x}(0), \boldsymbol{J_w}(k)\} = 0$$

$$var\{\boldsymbol{J_x}(0)\} = \boldsymbol{V}_{\widetilde{J_x}}$$

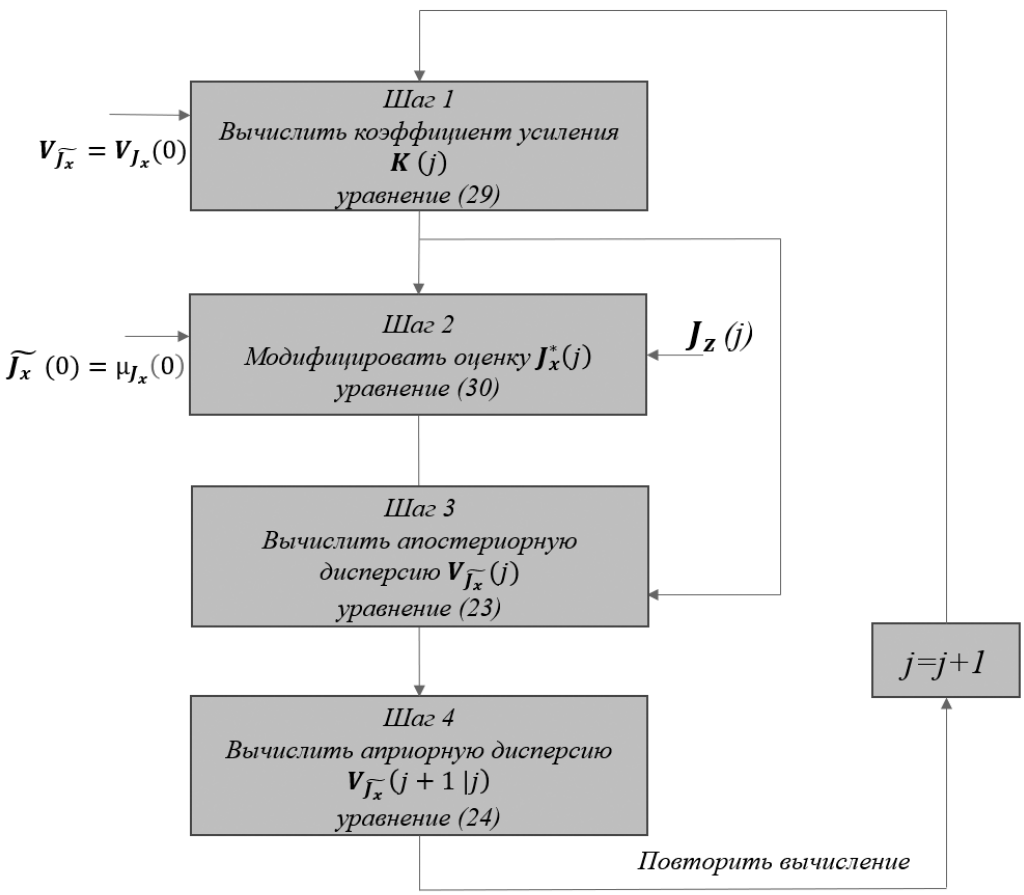

Рис. 1. Структурная схема вычислений по алгоритму фильтрации

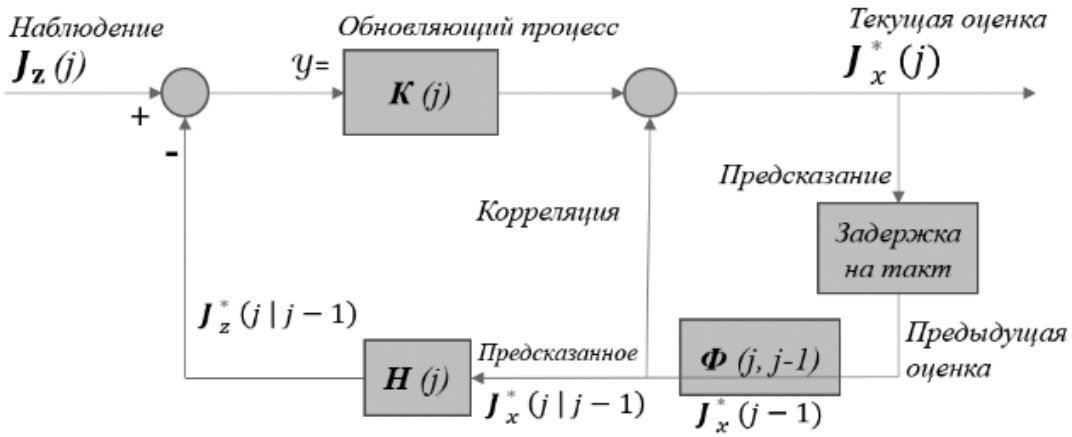

Рис. 2. Структурная схема дискретного фильтра

Алгоритм фильтрации:

$$\boldsymbol{J}_x^*(j) = \Phi(j, j-1)\boldsymbol{J}_x^*(j-1) + \\ + \boldsymbol{K}(j)\left[\boldsymbol{J}_z(j) - \boldsymbol{H}_{J_x}(j)\Phi(j, j-1)\boldsymbol{J}_x^*(j-1)\right] \quad (8)$$

Вычисление коэффициента усиления:

$$\boldsymbol{K}(j) = \boldsymbol{V}_{\widetilde{J_x}}(j|j-1)\boldsymbol{H}^T(j)$$

$$\left[\boldsymbol{H}(j)\boldsymbol{V}_{\widetilde{J_x}}(j|j-1)\boldsymbol{H}^T(j) + \boldsymbol{V}_{J_v}(j)\right]^{-1} = \quad (9)$$

$$= \boldsymbol{V}_{\widetilde{J_x}}(j)\boldsymbol{H}^T(j)\boldsymbol{V}_{J_v}{}^{-1}(j)$$

Величина априорной дисперсии:

$$\boldsymbol{V}_{\widetilde{J_x}}(j+1|j) = \Phi(\boldsymbol{j}+\boldsymbol{1}, j)\boldsymbol{V}_{\widetilde{J_x}}(j) \times \\ \times \Phi^T(\boldsymbol{j}+1, j) + \boldsymbol{\Gamma}(j)\boldsymbol{V}_{J_w}(j)\boldsymbol{\Gamma}^T(j) \quad (10)$$

Уравнение для апостериорной дисперсии:

$$\boldsymbol{V}_{\widetilde{J_x}}(j) = [\boldsymbol{I}_x - \boldsymbol{K}(j)\boldsymbol{H}(j)]\boldsymbol{V}_{\widetilde{J_x}}(j \mid j-1) \quad (11)$$

Начальные условия:

$$\boldsymbol{J}_x^*(0 \mid 0) = \boldsymbol{J}_x^*(0) = \\ = \mu_{J_x}(0), \boldsymbol{V}_{\widetilde{J_x}}(0|0) = \boldsymbol{V}_{\widetilde{J_x}}(0) = \boldsymbol{V}_{J_x} \quad (0)$$

Анализ структурой схемы рис. 1 показывает, что в фильтре реализуется идея предсказания – коррекции. Предыдущая оценка $\boldsymbol{J}_x^*(j-1)$ экстраполируется на один шаг вперед и затем используется для получения наилучшей оценки нового наблюдения $\boldsymbol{J}_z(j)$ основанной на всех предыдущих наблюдениях. Ошибка между «наилучшей оценкой» текущего наблюдения и фактическим наблюдением, а именно $\mathcal{Y}(j|j-1)$ или $\widetilde{J_z}(j \mid j-1)$, представляет собой новую информацию [компоненту $\boldsymbol{J}_z(j)$, ортогональную $\boldsymbol{J}_z(j-1)$].

Ошибка взвешивается с весом $\boldsymbol{K}(j)$, учитывающим значение дисперсий входного процесса, измерения и ошибки оценивания для формирования сигнала коррекции. Сигнал коррекции складывается с предсказанной оценкой и в результате получается новая оценка [2].

Структура дискретного фильтра, соответствующая уравнению (8) и изображённая на рис. 2, очень напоминает структуру исходной модели сообщения, заданной уравнением (6). Алгоритм фильтрации строится на использовании «обновляющей» компоненты, которая содержит новую информацию, полученную в результате наблюдения [2].

Результаты основных и производных параметров приведены в таблицах 1 и 3. Результаты средних значений основных и производных параметров сведены в таблицы 2 и 4.

Результаты исследований

Проанализируем параметры, полученные в некоторых системах искусственного интеллекта разных поколений. Для определения параметров применялся апробированный программный комплекс оценки информационных параметров нейролингвистической текстовой идентификации интеллектуальных систем [2], интерфейс которого приведен на рис. 3.

Исследованные семь систем искусственного интеллекта относятся к разным поколе-

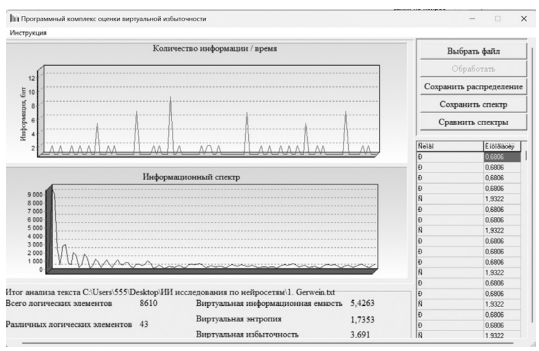

Рисунок 3. Интерфейс комплекса оценки параметров нейролингвистической текстовой идентификации систем искусственного интеллекта СИИ

Таблица 1

Параметры нейролингвистической текстовой идентификации систем искусственного интеллекта

Системы искусственного интеллекта	Информационная емкость, H_{Bmax}	Энтропия, H_B	Избыточность, B
СИИ-1 (Gerwin)	5,42	1,73	3,69
СИИ-2 (YandexGPT)	5,93	1,79	4,14
СИИ-3 (Micrasoft Copilot)	6,02	1,85	4,17
СИИ-4 (RoboGPT)	5,24	1,7	3,53
СИИ-5 (CopyMonkey)	5,42	1,74	3,67
СИИ-6 (Всезнайка)	5,67	1,93	3,74
СИИ-7 (ChatGPT 3.5)	3,8	1,65	2,14

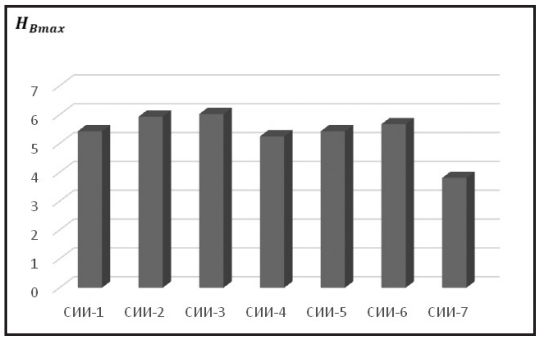

Рис. 4. Информационная емкость нейролингвистической текстовой информации систем ИИ

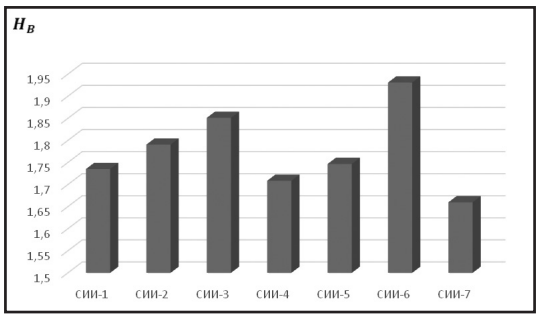

Рис. 5. Энтропия нейролингвистической текстовой информации систем ИИ

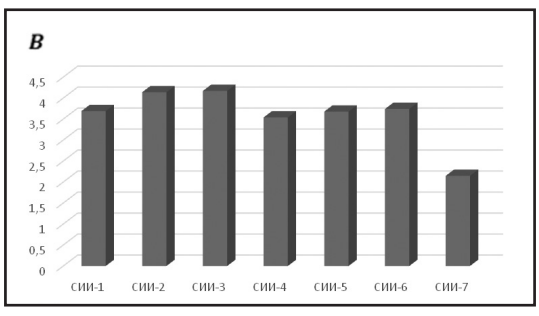

Рис. 6. Избыточность нейролингвистической текстовой информации систем ИИ

ниям и являются как российскими, так и иностранными. Результаты оценки основных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта сведены в таблицу 1.

Результаты исследования зависимости основных параметров нейролингвистической текстовой идентификации систем ИИ приведены на рисунках 4, 5 и 6.

Полученные результаты показывают, что минимальное значение информационной емкости текста систем искусственного интеллекта соответствует ChatGPT 3.5 – $\mathrm{H_{Bmax}} = 3{,}80$. Максимальное значение информационной емкости текста систем искусственного интеллекта соответствует Micrasoft Copilot – $\mathrm{H_{Bmax}} = 6{,}02$. Среднее значение информационной емкости – $\dot{\mathrm{H}}_{\mathrm{Bmax}} = 4{,}91$. Минимальное значение энтропии текста систем искусственного интеллекта соответствует ИС-Р1 – $\mathrm{H_B} = 1{,}57$. Максимальное значение энтропии текста систем искусственного интеллекта соответствует Micrasoft Copilot – $\mathrm{H_B} = 1{,}85$. Среднее значение энтропии – $\mathrm{H_B} = 1{,}71$. Минимальное значение избыточности текста систем искусственного интеллекта соответствует ChatGPT 3.5 – $\mathrm{B} = 2{,}14$. Максимальное значение избыточности текста систем искусственного интеллекта соответствует Micrasoft Copilot – $\mathrm{B} = 4{,}17$. Среднее значение виртуальной избыточности – $\mathrm{B} = 3{,}15$.

Средние значения основных параметров нейролингвистических текстовых идентификаторов систем искусственного интеллекта сведены в таблицу 2.

В таблицу 3 сведены результаты расчётов производных параметров нейролингвистической текстовой идентификации систем ис-

Таблица 2

Среднее значение параметров нейролингвистических текстовых идентификаторов систем искусственного интеллекта

Среднее значение информационной емкости, H_{Bmax}	Среднее значение энтропии, H_B	Среднее значение избыточности, B
4,91	1,71	3,15

Таблица 3

Производные параметры систем искусственного интеллекта

СИИ	Коэффициент избыточности, μ_B	Коэффициент вербальности, G_B
СИИ-1 (Gerwin)	0,68	0,47
СИИ-2 (YandexGPT)	0,698	0,432
СИИ-3 (Micrasoft Copilot)	0,692	0,443
СИИ-4 (RoboGPT)	0,674	0,482
СИИ-5 (CopyMonkey)	0,678	0,474
СИИ-6 (Всезнайка)	0,659	0,516
СИИ-7 (ChatGPT 3.5)	0,561	0,772

кусственного интеллекта: коэффициент избыточности и коэффициент вербальности.

Результаты исследования зависимости производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта приведены на рисунках 7 и 8.

Установлено, что минимальное значение коэффициента избыточности текста систем искусственного интеллекта соответствует ChatGPT 3.5 – μ_B = 0,561. Максимальное значение коэффициента избыточности текста систем искусственного интеллекта соответствует YandexGPT – μ_B = 0,6980. Среднее значение коэффициента избыточности – μ_B = 0,629. Минимальное значение коэффициента вербальности текста систем искусственного интеллекта соответствует YandexGPT – G_B = 0,432. Максимальное значение коэффициента вербальности текста систем искусственного интеллекта соответствует ChatGPT 3.5 – G_B = 0,772. Среднее значение коэффициента вербальности – G_B = 0,602.

Средние значения коэффициентов избыточности и вербальности систем искусственного интеллекта сведены в таблицу 4.

Исследования, проведённые авторами, показали, что проблема идентификации мозговых механизмов речевой деятельности может быть решена при использовании в качестве идентификатора количества информации, содержащегося в логических элементах речи (словах, слогах и т.п.).

В качестве параметров идентификатора в данном случае выступают: коэффициент избыточности, коэффициент вербальности, эм-

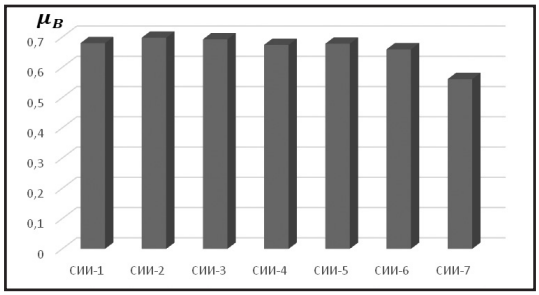

Рис. 7. Коэффициент избыточности нейролингвистической текстовой информации систем ИИ

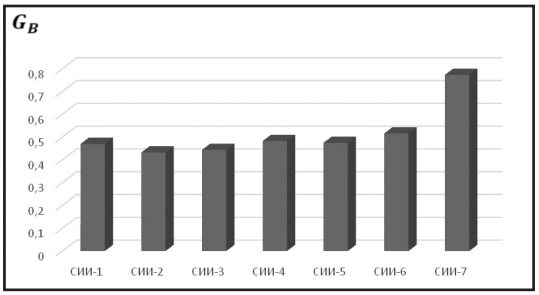

Рис. 8. Коэффициенты вербальности нейролингвистической текстовой информации систем ИИ

Таблица 4

Среднее значение коэффициентов избыточности и вербальности систем искусственного интеллекта

Среднее значение коэффициента избыточности, μ_B	Среднее значение коэффициента вербальности, G_B
0,629	0,602

пирическая энтропия, информационная емкость. Все поставленные изначально задачи решены. Определено, что значения основных и производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта при постановке одной и той же задачи разным системам ИИ значительно отличаются. Проанализированы основные и производные параметры нейролингвистической текстовой идентификации систем искусственного интеллекта. Определены средние значения основных и производных параметров нейролингвистических текстовых идентификаторов систем искусственного интеллекта.

Заключение

В исследовании проанализировано семь систем искусственного интеллекта: одна система третьего поколения СИИ-7 (ChatGPT 3.5), шесть систем четвертого поколения (Gerwin, YandexGPT, RoboGPT, CopyMonkey, Micrasoft Copilot, Всезнайка). Всем системам ИИ определена одна и та же задача для решения. Постановка задачи была на русском языке. Исследование показало, что результат каждой СИИ индивидуален и отличается от всех остальных. Таким образом наблюдается, что показатели параметров систем ИИ третьего поколения значительно отличается от показателей параметров систем ИИ четвертого поколения. При переходе от одной системы ИИ к другой наблюдается изменение параметров. В среднем это выражается в изменении на 40 %: информационная емкость – на 36 %; энтропия на 15 %; избыточность на 48 %; коэффициента избыточности на 20 %; коэффициент вербальности на 44%.

В результате проведенных исследований установлено следующее. Во-первых, что значения основных и производных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта при постановке одной и той же задачи разным системам искусственного интеллекта значительно отличаются. Во-вторых, информационные характеристики текстовой составляющей систем искусственного интеллекта можно использовать в качестве параметров нейролингвистической текстовой идентификации систем искусственного интеллекта. В-третьих, в качестве преимущественных параметров нейролингвистической текстовой идентификации систем искусственного интеллекта целесообразно использовать из основных параметров информационную емкость и избыточность, из производных параметров коэффициент вербальности, поскольку наблюдается максимальное изменение показателей данных параметров при переходе от одной системы искусственного интеллекта к другой (в среднем на 40 %).

Основной проблемой современности является повсеместное применения ИИ в жизни человека, невозможно определить решена та или иная задача с помощью систем ИИ или самостоятельно человеком, такая закономерность настораживает. Такие неблагоприятные перспективы приводят к необходимости в глубоком изучении данного вопроса и поиску пути его решения. Полученные результаты открывают возможность новых подходов к обнаружению идентификации и аутентификации систем искусственного интеллекта. Исследования данной проблемы позволяют по параметрам нейролингвистической текстовой идентификации определить, использовались системы искусственного интеллекта для решения поставленной задачи или задача решалась человеком. Это обеспечит информационную безопасность передаваемых данных в телекоммуникационных системах и сетях. Полученные результаты позволяют показать индивидуальность систем искусственного интеллекта, откуда следует возможность существования в них задатков моделирования мыслительной деятельности.

Литература

1. Котенко В.В. Технологии информационного анализа пользовательского уровня телекоммуникационных систем. Учебное пособие // Ростов-на-Дону – Таганрог: Издательство Южного федерального университета, 2019. – 194 с

2. Котенко В.В., Румянцев К.Е. Теория информации. Учебное пособие // Ростов-на-Дону – Таганрог: Издательство Южного федерального университета, 2018. – 239 с.

3. Котенко В.В. Теория виртуализации и защита телекоммуникаций: монография – Таганрог: Изд-во ТТИ ЮФУ, 2011. – 236 с.

4. Котенко В.В. Румянцев К.Е., Котенко С.В. Методология идентификационного анализа инфокоммуникационных систем. Монография. – Ростов–на–Дону: Издательство Южного федерального университета, 2014. – 315 с.

5. Котенко В.В., Румянцев К.Е. Теоретические основы противодействия угрозам терроризма: монография / Издательство: Южный федеральный университет (Ростов-на-Дону) 2014 г. – 228 с.

6. Хаджиева Л. К. Анализ аутентификации идентификаторов уровня систем разграничения доступа. // Вестник науки. Сборник научных статей по материалам Международной научно-практической конференции «Перспективы развития науки в современном мир». Уфа, 7 апреля 2023. Часть 1. – Уфа: Издательство «НИЦ Вестник науки», 2023. – С. 19-27.

7. Котенко В.В., Хаджиева Л.К. Аутентификация и системы разграниченного доступа пользовательского уровня как меры защиты информации в банках // Сборник статей XII Всероссийской научно-практической конференции «Молодежь, наука, инновации», Грозный, 18 октября 2023 г. – С – 72-76.

8. Litvinova T. Authorship attribution of russian social media texts: does the volume of data favor idiolect identification? / Lecture Notes in Networks and Systems. – 2022. – T. 315 LNNS. – P. 352-362.

9. Symmetric Multi-Scale Residual Network Ensemble with Weighted Evidence Fusion Strategy for Facial Expression Recognition. / Juan Liu, Min Hu, Ying Wang, Zhong Huang, Julang Jiang // Symmetry. – 2023. – 15 (6). – P. 1228. – https://doi.org/10.3390/sym15061228

10. Ямченко Ю.В. Методы решения задач аутентификации и идентификации пользователя на основе анализа клавиатурного почерка // Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия Приборостроение. – 2020. – № 1 (130). – С. 124-139.

11. Куртукова, А. В. Разработка интеллектуальной системы для идентификации автора исходного кода на основе нейронных сетей / А. В. Куртукова. — Текст: непосредственный // Молодой ученый. — 2020. — № 40 (330). — С. 2-3.

12. Куртукова А.В., Романов А.С., Соболев А.А. Воздействие корпоративных стандартов разработки на идентификацию автора исходного кода // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. 2020. № 1-2. С. 140-142.

13. Куртукова А.В., Романов А.С. Проблема искусственно сгенерированных исходных кодов в задаче идентификации автора программы // Сборник избранных статей научной сессии ТУСУР. 2022. № 1-2. С. 131-134.

14. Куртукова А.В., Романов А.С., Федотова А.М., Шелупанов А.А. Архитектура интеллектуальной системы для идентификации автора исходного кода// Доклады Томского государственного университета систем управления и радиоэлектроники. 2022. Т. 25. № 3. С. 39-44.

15. Кетермина Т.С., Тагиров Т.М. Элементы искусственного интеллекта в решении задач анализа текстов // Computational nanotechnology. 2022 N. 9. №2. С. 35-44. DOI: 10. 33693/2313-223X-2022-9-2-35-44.

16. Черкасов А.Н., Туркин Е.А. Выбор оптимальной архитектуры искусственной нейронной сети для задачи классификации текстов // Ежеквартальный рецензируемый, реферируемый научный журнал «Вестник АГУ». Выпуск 1 (276) 2021. С. 62-66

17. Уздяев М.Ю. Распознавание агрессивных действий с использованием нейросетевых архитектур 3d-cnn. // Известия ТулГУ. Технические науки. 2020. Выпуск 2 С. 316- 329

18. Батраева И.А., Нарцев А. Д., Лезгян А.С. Использование анализа семантической сложности слов при определении задач определения жанровых принадлежностей текстов, методов глубокого обучения. // Вестник томского государственного университета № 50 2020. С. 14-21

19. Частикова В.А., Казачев К.В., Гуляй В.Г. Методы обработки естественного языка в решении задач обнаружения атак социальной инженерии. // Ежеквартальный рецензируемый, реферируемый научный журнал «Вестник АГУ». Выпуск 1 (291) 2021. С. 95-107

20. Глазкова А.В. Сравнение нейросетевых моделей для классификации текстовых фрагментов, содержащих биографическую информацию. // Тюменский государственный университет. Т. 32. №2 С. г. Тюмень, 625003, Россия 2019 С. 65-69

21. https://ru.wikipedia.org/wiki/Нейролингвистика

22. https://smart-estet.ru/articles/sistemy-iskusstvennogo-intellekta

References

1. Kotenko V.V. Tekhnologii informatsionnogo analiza pol'zovatel'skogo urovnya telekommunikatsionnykh sistem. Uchebnoye posobiye // Rostov-na-Donu – Taganrog: Izdatel'stvo Yuzhnogo federal'nogo universiteta, 2019. – 194 s

2. Kotenko V.V., Rumyantsev K.Ye. Teoriya informatsii. Uchebnoye posobiye // Rostov-na-Donu – Taganrog: Izdatel'stvo Yuzhnogo federal'nogo universiteta, 2018. – 239 s.

3. Kotenko V.V. Teoriya virtualizatsii i zashchita telekommunikatsiy: monografiya – Taganrog: Izd-vo TTI YUFU, 2011. – 236 s.

4. Kotenko V.V. Rumyantsev K.Ye., Kotenko S.V. Metodologiya identifikatsionnogo analiza infokommunikatsionnykh sistem. Monografiya. – Rostov–na–Donu: Izdatel'stvo Yuzhnogo federal'nogo universiteta, 2014. – 315 s.

5. Kotenko V.V., Rumyantsev K.Ye. Teoreticheskiye osnovy protivodeystviya ugrozam terrorizma: monografiya / Izdatel'stvo: Yuzhnyy federal'nyy universitet (Rostov-na-Donu) 2014 g. – 228 s.

6. Khadzhiyeva L. K. Analiz autentifikatsii identifikatorov urovnya sistem razgranicheniya dostupa. // Vestnik nauki. Sbornik nauchnykh statey po materialam Mezhdunarodnoy nauchno-prakticheskoy konferentsii «Perspektivy razvitiya nauki v sovremennom mir». Ufa, 7 aprelya 2023. Chast' 1. – Ufa: Izdatel'stvo «NITS Vestnik nauki», 2023. – S. 19-27.

7. Kotenko V.V., Khadzhiyeva L.K. Autentifikatsiya i sistemy razgranichennogo dostupa pol'zovatel'skogo urovnya kak mery zashchity informatsii v bankakh // Sbornik statey XII Vserossiyskoy nauchno-prakticheskoy konferentsii «Molodezh', nauka, innovatsii», Groznyy, 18 oktyabrya 2023 g. – S – 72-76.

8. Litvinova T. Authorship attribution of Russian social media texts: does the volume of data favor idiolect identification? / Lecture Notes in Networks and Systems. – 2022. – Vol. 315 LNNS. – P. 352-362.

9. Symmetric Multi-Scale Residual Network Ensemble with Weighted Evidence Fusion Strategy for Facial Expression Recognition. / Juan Liu, Min Hu, Ying Wang, Zhong Huang, Julang Jiang // Symmetry. – 2023. – 15 (6). – P. 1228. – https://doi.org/10.3390/sym15061228

10. Yamchenko YU.V. Metody resheniya zadach autentifikatsii i identifikatsii pol'zovatelya na osnove analiza klaviaturnogo pocherka // Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. N.E. Baumana. Seriya Priborostroyeniye. – 2020. – № 1 (130). – S. 124-139.

11. Kurtukova, A. V. Razrabotka intellektual'noy sistemy dlya identifikatsii avtora iskhodnogo koda na osnove neyronnykh setey / A. V. Kurtukova. — Tekst: neposredstvennyy // Molodoy uchenyy. — 2020. — № 40 (330). — S. 2-3.

12. Kurtukova A.V., Romanov A.S., Sobolev A.A. Vozdeystviye korporativnykh standartov razrabotki na identifikatsiyu avtora iskhodnogo koda // Elektronnyye sredstva i sistemy upravleniya. Materialy dokladov Mezhdunarodnoy nauchno-prakticheskoy konferentsii. 2020. № 1-2. S. 140-142.

13. Kurtukova A.V., Romanov A.S. Problema iskusstvenno sgenerirovannykh iskhodnykh kodov v zadache identifikatsii avtora programmy // Sbornik izbrannykh statey nauchnoy sessii TUSUR. 2022. № 1-2. S. 131-134.

14. Kurtukova A.V., Romanov A.S., Fedotova A.M., Shelupanov A.A. Arkhitektura intellektual'noy sistemy dlya identifikatsii avtora iskhodnogo koda// Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki. 2022. T. 25. № 3. S. 39-44.

15. Ketermina T.S., Tagirov T.M. Elementy iskusstvennogo intellekta v reshenii zadach analiza tekstov // Computational nanotechnology. 2022 N. 9. №2. S. 35-44. DOI: 10. 33693/2313-223KH-2022-9-2-35-44.

16. Cherkasov A.N., Turkin Ye.A. Vybor optimal'noy arkhitektury iskusstvennoy neyronnoy seti dlya zadachi klassifikatsii tekstov // Yezhekvartal'nyy retsenziruyemyy, referiruyemyy nauchnyy zhurnal «Vestnik AGU». Vypusk 1 (276) 2021. S. 62-66

17. Uzdyayev M.YU. Raspoznavaniye agressivnykh deystviy s ispol'zovaniyem neyrosetevykh arkhitektur 3d-cnn. // Izvestiya TulGU. Tekhnicheskiye nauki. 2020. Vypusk 2 S. 316- 329

18. Batrayeva I.A., Nartsev A. D., Lezgyan A.S. Ispol'zovaniye analiza semanticheskoy slozhnosti slov pri opredelenii zadach opredeleniya zhanrovykh prinadlezhnostey tekstov, metodov glubokogo obucheniya. // Vestnik tomskogo gosudarstvennogo universiteta № 50 2020. S. 14-21

19. Chastikova V.A., Kazachev K.V., Gulyay V.G. Metody obrabotki yestestvennogo yazyka v reshenii zadach obnaruzheniya atak sotsial'noy inzhenerii. // Yezhekvartal'nyy retsenziruyemyy, referiruyemyy nauchnyy zhurnal «Vestnik AGU». Vypusk 1 (291) 2021. S. 95-107

20. Glazkova A.V. Sravneniye neyrosetevykh modeley dlya klassifikatsii tekstovykh fragmentov, soderzhashchikh biograficheskuyu informatsiyu. // Tyumenskiy gosudarstvennyy universitet. T. 32. №2 S. g. Tyumen', 625003, Rossiya 2019 S. 65-69.

21. https://ru.wikipedia.org/wiki/Нейролингвистика

22. https://smart-estet.ru/articles/sistemy-iskusstvennogo-intellekta

КОТЕНКО Владимир Владимирович, кандидат технических наук, доцент кафедры «Информационная безопасность телекоммуникационных систем». Институт компьютерных технологий и информационной безопасности ЮФУ. 347922, Ростовская область, г. Таганрог, ул. Чехова, 2. E-mail: kotenkovv@sfedu.ru

РУМЯНЦЕВ Константин Евгеньевич, доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность телекоммуникационных систем». Институт компьютерных технологий и информационной безопасности ЮФУ. 347922, Ростовская область, г. Таганрог, ул. Чехова, 2. E-mail: rumyancev@sfedu.ru

ХАДЖИЕВА Лаура Куйраевна, соискатель кафедры «Информационная безопасность телекоммуникационных систем». Институт компьютерных технологий и информационной безопасности ЮФУ. 347922, Ростовская область, г. Таганрог, ул. Чехова, 2.; старший преподаватель кафедры «Информатика и вычислительная техника». Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова. 364061, Чеченская Республика., г. Грозный, пр. Х. Исаева, 100. E-mail: laura.hadjieva3009@mail.ru

KOTENKO Vladimir Vladimirovich, Candidate of Technical Sciences, Associate Professor of the «Information Security of Telecommunication Systems». Institute of Computer Technologies and Information Security of SFU. 347922, Rostov Region, Taganrog, Chekhov St., 2. E-mail: kotenkovv@sfedu.ru

RUMYANTSEV Konstantin Evgenievich, Doctor of Technical Sciences, Professor, Head of the Department «Information Security of Telecommunication Systems». Institute of Computer Technologies and Information Security of SFU. 347922, Rostov Region, Taganrog, Chekhov St., 2. E-mail: rumyancev@sfedu.ru

KHADZHIEVA Laura Kuiraevna, applicant of the Department of «Information Security of Telecommunication Systems». Institute of Computer Technologies and Information Security of SFU. 347922, Rostov Region, Taganrog, Chekhov St., 2.; senior lecturer at the Department of Computer Science and Computer Engineering. Grozny State Petroleum Technical University named after Academician M.D. Millionshchikov. 364061, Chechen Republic., Grozny, Kh. Isaev Ave., 100. E-mail: laura.hadjieva3009@mail.ru

УДК 004.056 Вестник УрФО № 2(52) / 2024, с. 24–36

Хорев А. А., Порсев И. С.

DOI: 10.14529/secur240203

ВЕРОЯТНОСТНЫЙ МЕТОД ОБОСНОВАНИЯ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ ЭФФЕКТИВНОСТИ ЗАЩИТЫ РЕЧЕВОЙ ИНФОРМАЦИИ ОТ ЕЕ УТЕЧКИ ПО ТЕХНИЧЕСКИМ КАНАЛАМ

В статье предложено в качестве показателей оценки эффективности защиты речевой информации использовать вероятности вскрытия тематики перехваченного разговора и вероятности составления его аннотации. Получены аналитические соотношения по расчету вероятностей вскрытия тематики перехваченного разговора и составления его аннотации в зависимости от словесной и фразовой разборчивости речи, а также от количества ключевых слов и ключевых фраз, необходимых для определения тематики разговора и составления его аннотации. Предложена методика обоснования критериев эффективности защиты речевой информации.

Приведены результаты оценки зашумленных речевых сигналов методом артикуляционных испытаний. Получены аналитические соотношения для расчета словесной и фразовой разборчивости речи в зависимости от отношения сигнал/шум для 4-х видов шумов: «белый» шум, «розовый» шум», «коричневый» шум» и шумовая «речеподобная» помеха.

Для сравнительной оценки эффективности шумовых помех введено понятие «коэффициент эффективности шумовой помехи», под которым понимается отношение сигнал/шум для определенного вида шума к отношению сигнал/шум для «белого» шума, выраженных в дБ, при которой разборчивости речи одинаковы. Получены аналитические соотношения для расчета «коэффициент эффективности шумовой помехи» для трех видов шумовых помех.

Ключевые слова: *защита акустической речевой информации, показатели и критерии оценки эффективности защиты речевой информации от ее утечки по техническим каналам, разборчивость речи, вероятностный метод обоснования показателей и критериев оценки эффективности защиты речевой информации.*

Horev A. A., Porsev I. S.

A PROBABILISTIC METHOD FOR SUBSTANTIATING INDICATORS AND CRITERIA FOR THE EFFECTIVENESS OF PROTECTING SPEECH INFORMATION FROM ITS LEAKAGE THROUGH TECHNICAL CHANNELS

The article suggests using probabilities of revealing the subject of an intercepted conversation and the probability of compiling its annotation as indicators for evaluating the effectiveness of protecting speech information. Analytical ratios have been obtained for calculating the probabilities of opening the subject of an intercepted conversation and compiling its annotation, depending on the verbal and phrasal intelligibility of speech, as well as on the number of keywords and key phrases necessary to determine the topic of the conversation and compile its annotation. A methodology for substantiating the criteria for the effectiveness of speech information protection is proposed.

The results of the evaluation of noisy speech signals by the method of articulation tests are presented. Analytical ratios have been obtained for calculating verbal and phrasal intelligibility of speech depending on the signal-to-noise ratio for 4 types of noise: "white" noise, "pink" noise", "brown" noise" and noise "speech-like" interference.

For a comparative assessment of the effectiveness of noise interference, the concept of "noise interference efficiency coefficient" is introduced, which refers to the signal-to-noise ratio for a certain type of noise to the signal-to-noise ratio for "white" noise, expressed in dB, at which speech intelligibility is the same. Analytical relations have been obtained for calculating the "noise interference efficiency coefficient" for three types of noise interference.

Keywords: *protection of acoustic speech information, indicators and criteria for evaluating the effectiveness of protecting speech information from leakage through technical channels, speech intelligibility, probabilistic method for substantiating indicators and criteria for evaluating the effectiveness of speech information protection.*

Введение

В качестве показателя эффективности защиты речевой информации от ее утечки по техническим каналам наиболее часто используется разборчивость речи (W), отображающая качественную область понятности перехваченного разговора, под которой понимается отношение количества правильно распознанных слов (фраз) к общему количеству слов (фраз) в перехваченном разговоре.

В качестве критерия эффективности защиты речевой информации от ее утечки по техническим каналам используется пороговый критерий, при использовании которого считается, что принятые меры по защите речевой информации от ее утечки по техническим каналам эффективны, если разборчивость речи, перехваченная средствами разведки по данным каналам (W), не превышает установленного порогового значения ($W_п$).

Таблица 1.

Показатели и критерии эффективности защиты речевой информации от ее утечки по техническим каналам

Цели защиты речевой информации	Показатели эффективности защиты речевой информации	Критерии эффективности защиты речевой информации	Пороговое значение разборчивости речи
Скрытие тематики текста	Словесная разборчивость текста ($W_{сл}$)	Количество правильно распознанных слов не позволяет установить тематику перехваченного разговора ($W_{сл} \le W_{сл.п}$)	$W_{сл.п} = 0{,}2$
Скрытие содержания текста	Словесная разборчивость текста ($W_{сл}$)	Количество правильно распознанных слов не позволяет составить аннотацию (краткую справку) о перехваченном разговоре ($W_{сл} \le W_{сл.п}$)	$W_{сл.п} = 0{,}3$

Критерии эффективности защиты речевой информации во многом зависят от целей, преследуемых при защите информации, например, скрыть тематику ведущегося разговора или скрыть его смысловое содержание.

В работе [1] на основе метода экспертных оценок предложены следующие критерии эффективности защиты речевой информации от ее утечки по техническим каналам, которые представлены в таблице 1.

Однако не все распознанные слова или фразы относятся к ключевым, по которым можно установить тематику перехваченного разговора и составить его аннотацию.

Следовательно, для оценки эффективности защиты речевой информации необходимо не только рассчитать словесную или фразовую разборчивость речи, но и оценить количество ключевых слов и фраз, требуемых для определения тематики разговора или составления ее аннотации.

1. Вероятностный метод обоснования критериев эффективности защиты речевой информации от ее утечки по техническим каналам

Будем полагать, что тематика разговора вскрыта, если количество распознанных ключевых слов $N_{сл.кл.р}$ будет не менее установленного порогового значения $N_{сл.кл.min}$, а аннотация разговора составлена, если количество распознанных ключевых фраз $N_{фр.кл.р}$ будет не менее установленного порогового значения $N_{фр.кл.min}$.

Тогда, вероятность вскрытия тематики разговора P_m можно рассчитать по формулам [2]:

$$P_m(N_{сл.кл.р} \ge N_{сл.кл.min}) = \sum_{i=N_{сл.кл.min}}^{N_{сл.кл.р}} P_m(N_{сл.кл.р.i}) \quad (1)$$

$$P_m(N_{сл.кл.р.i}) = \frac{C_{N_{сл.кл.}}^{N_{сл.кл.р.i}} \cdot C_{N_{сл.}-N_{сл.кл.}}^{N_{сл.кл}-N_{сл.кл.р.i}}}{C_{N_{сл.}}^{N_{сл.кл}}} \quad (2)$$

где $C_b^a = \frac{b!}{a! \cdot (b-a)!}$ – формула комбинаторики, определяющая количество сочетаний без повторений [3];

$N_{сл}$ – количество слов в перехваченном разговоре;

$N_{сл.кл} = k_{сл} \times N_{сл}$ – количество ключевых слов в перехваченном разговоре;

$k_{сл}$ – среднее относительное количество ключевых слов по данной тематике;

$N_{сл.р} = W_{сл} \times N_{сл}$ – количество распознанных слов в перехваченном разговоре;

$W_{сл}$ – словесная разборчивости перехваченного разговора;

$N_{сл.кл.р} = W_{сл} \times N_{сл.кл} = W_{сл} \times k_{сл} \times N_{сл}$ – количество распознанных ключевых слов в перехваченном разговоре;

$N_{сл.кл.min}$ – минимальное количество ключевых слов, необходимых для вскрытия тематики перехваченного разговора.

Учитывая, что $b > a$ формулу комбинаторики запишем в виде:

$$C_b^a = \frac{b!}{a!\cdot(b-a)!} = \frac{(a+1)\cdot(a+2)\cdot(a+3)\cdots b}{(b-a)!} = \frac{\prod_{j=a+1}^{b} j}{\prod_{j=1}^{b-a} j} \quad (3)$$

Подставив (3) в (2) и сократив формулу получим:

$$P_m(N_{сл.кл.p.i}) = \frac{\prod_{j=N_{сл.кл.p.i}+1}^{N_{сл.кл.}} j \cdot \prod_{j=N_{сл.кл.}-N_{сл.кл.p.i}+1}^{N_{сл..}-N_{сл.кл.}} j \cdot \prod_{j=1}^{N_{сл..}-N_{сл.кл.}} j}{\prod_{j=1}^{N_{сл.кл.}-N_{сл.кл.p.i}} j \cdot \prod_{j=N_{сл.кл.}-N_{сл.кл.p.i}+1}^{(N_{сл..}-N_{сл.кл.})-(N_{сл.кл.}-N_{сл.кл.p.i.})} j \cdot \prod_{j=N_{сл.кл.}+1}^{N_{сл..}} j} \quad (4)$$

Аналогичным образом может быть рассчитана и вероятность составления аннотации разговора $P_{ан}$:

$$P_{ан}(N_{фр.кл.p} \geq N_{фр.кл.min}) = \sum_{i=N_{фр.кл.min}}^{N_{фр.кл.p}} P_{ан}(N_{фр.кл.p.i}) \quad (5)$$

$$P_{ан}(N_{фр.кл.p.i}) = \frac{\prod_{j=N_{фр.кл.p.i}+1}^{N_{фр.кл.}} j \cdot \prod_{j=N_{фр.кл.}-N_{фр.кл.p.i}+1}^{N_{фр..}-N_{фр.кл.}} j \cdot \prod_{j=1}^{N_{фр..}-N_{фр.кл.}} j}{\prod_{j=1}^{N_{фр.кл.}-N_{фр.кл.p.i}} j \cdot \prod_{j=N_{фр.кл.}-N_{фр.кл.p.i}+1}^{(N_{фр}-N_{фр.кл.})-(N_{фр.кл.}-N_{фр.кл.p.i.})} j \cdot \prod_{j=N_{фр.кл.}+1}^{N_{фр..}} j} \quad (6)$$

где $N_{фр}$ – количество фраз в перехваченном разговоре;

$N_{фр.кл}$ – количество ключевых фраз в перехваченном разговоре;

$N_{фр.кл.p}$ – количество распознанных ключевых слов в перехваченном разговоре;

$N_{фр.кл.min}$ – минимальное количество ключевых фраз, необходимых для вскрытия тематики перехваченного разговора;

$W_{фр}$ – словесная разборчивости перехваченного разговора.

2. Экспериментальные исследования разборчивости речи условиях шумов

С целью оценки разборчивости речи в условиях различного вида шумов были проведены экспериментальные исследования.

Исследования проводились в четыре этапа.

На **первом этапе** формировались исходные записи тестовых речевых сигналов.

В качестве тестовых речевых сигналов использовались таблицы слов из ГОСТ 16600-72 «Требования к разборчивости речи и методы артикуляционных измерений» и фраз – из ГОСТ 50840-95 «Передача речи по трактам связи методы оценки качества, разборчивости и узнаваемости». Запись тестовых речевых сигналов проводили 4 диктора (2 мужчин и 2 женщины), которые не имели дефектов речи. Чтение слов дикторами осуществлялось ровным голосом, четко и без подчеркивания отдельных звуков с постоянным уровнем речи. Дикторы выдерживали постоянный ритм речи на протяжении чтения всей таблицы.

Запись тестовых речевых сигналов проводилась в служебном помещении при уровне шума не более 35–40 дБ на профессиональный цифровой диктофон Tascam DR-05X. Режимы записи: 48 кГц, 24 бит, формат записи – wav. Запись проводилась на съемную SD – карту.

Всего каждым диктором было записано 4 аудиофайла: 2 аудиофайла

с 2 таблицами по 50 слов и 2 аудиофайла с 2 таблицами по 50 фраз.

На **втором этапе** формировались помеховые сигналы.

При проведении экспериментальных исследований использовались следующие наиболее часто используемых в системах активной защиты видов шумов:

1) «белый» шум (БШ) – шум с равномерной спектральной плотностью;

2) «розовый» шум» (РШ) – шум со спадом спектральной плотности

3) 3 дБ/на октаву;

4) «коричневый» шум» (КШ) – шум со спадом спектральной плотности 6 дБ/на октаву;

5) шумовая «речеподобная» помеха (ШРП) – шум с огибающей спектра, подобной огибающей спектра речевого сигнала.

Таблица 2.

Типовые уровни речевого сигнала в октавных полосах частотного диапазона речи L_{si}

Номер полосы	Частотные границы полосы ($f_н$ - $f_в$), Гц	Среднегеометрическая частота полосы f_i, Гц	Типовые спектральные уровни речи L_{si}, измеренные на расстоянии 1 м от источника сигнала, дБ			
			Тихая речь (64 дБ)	Речь средней громкости (70 дБ)	Громкая речь (76 дБ)	Очень громкая речь (84 дБ)
1	90 - 175	125	47	53	59	67
2	175 - 355	250	60	66	72	80
3	355 - 710	500	60	66	72	80
4	710 - 1400	1000	55	61	67	75
5	1400 - 2800	2000	50	56	62	70
6	2800 - 5600	4000	47	53	59	67
7	5600 - 11200	8000	43	49	55	63

Для формирования различного вида шумов использовалась ПЭВМ и специальное программное обеспечение Adobe Audition CC 2019 (далее – ПО Audition).

Для формирования шумовой «речеподобной» помехи использовался типовой спектр речевого сигнала, приведенный в таблице 2 [5].

На **третьем этапе** формировались тестовые речевые сигналы с заданным отношением сигнал/шум.

При проведении экспериментальных исследований под отношением сигнал/шум (с/ш) понималось среднее отношение речевой с/ш за весь период чтения диктором артикуляционной таблицы (50 слов или 50 фраз).

Для измерения среднего уровня тестового речевого сигнала из аудиофайла с записью таблиц слов или таблиц фраз, воспроизводимых диктором, с помощью ПО Audition вырезались паузы между словами.

Аудиофайл с записью тестового речевого сигнала с вырезанными паузами воспроизводился через звуковую колонку и с помощью шумомера, установленного на расстоянии 1 м от звуковой колонки, проводилось измерение его уровня.

С использованием инструмента «Усиление/Затухание» в ПО Audition устанавливался уровень усиления, при котором средний уровень тестового речевого сигнала с вырезанными паузами соответствовал уровню речевого сигнала средней громкости (см. таблицу 2).

Далее при фиксированном коэффициенте усиления ПО Audition аудиофайл с записью тестового речевого сигнала с паузами воспроизводился через звуковую колонку и осуществлялась его запись на цифровой диктофон Tascam DR-05X, установленный на расстоянии 1 м от звуковой колонки.

Затем на звуковую колонку с компьютера подавался шумовой сигнал и

с использованием инструмента «Усиление/Затухание» в ПО Audition устанавливался уровень шумового сигнала, при котором обеспечивалось требуемое отношение сигнал/шум частотном интервале 90–11200 Гц: 10 дБ;

5 дБ; 0 дБ; –5 дБ, –10 дБ, –15 дБ; –20 дБ.

Излучаемый звуковой колонкой шумовой сигнал записывался на цифровой диктофон Tascam DR-05X, установленный на расстоянии 1 м от звуковой колонки.

Аудиофайл с зашумленным тестовым речевым сигналом и заданным отношением с/ш формировался путем микширования тестового речевого сигнала и шумового сигнала с требуемым уровнем шума.

Всего каждого вида шумового сигнала было сформировано по 56 аудиофайлов с записями таблиц из 50 слов и по 56 аудиофайлов с записями таблиц по 50 фраз.

На **четвертом этапе** зашумленные тестовые речевые сигналы (аудиофайлы) прослушивались аудиторами и определялась словесная и фразовая разборчивость речи.

Прослушивание осуществляли 14 аудиторами (8 мужчин и 6 женщин), которые не имели дефектов слуха.

Таблица 3.

Результаты экспериментальных исследований

Отношение с/ш, дБ	Разборчивость речи, *W*							
	«Белый» шум»		«Розовый» шум		«Коричневый» шум		Шумовая «речеподобная помеха»	
	$W_{сл}$	$W_{фр}$	$W_{сл}$	$W_{фр}$	$W_{сл}$	$W_{фр}$	$W_{сл}$	$W_{фр}$
10 дБ	1	1	1	1	1	1	1	1
5 дБ	0,97	1	1	1	1	1	0,91	1
0 дБ	0,91	0,99	0,9	1	0,9	1	0,78	0,92
- 5 дБ	0,79	0,96	0,69	0,83	0,88	1	0,6	0,78
- 10 дБ	0,57	0,82	0,44	0,59	0,82	0,97	0,29	0,43
- 15 дБ	0,31	0,48	0,23	0,15	0,53	0,83	0,07	0,11
- 20 дБ	0,13	0,13	0	0	0,39	0,59	0	0

Испытания проводились в помещении в нормальных климатических условиях при уровне шума не более 35–40 дБ.

Записи прослушивались аудиторами через головные наушники, подключаемые к линейному выходу звуковой карты ПЭВМ и обеспечивающие высокий уровень изоляции. Услышанные слова и фразы аудиторами записывались в соответствующие таблицы.

Слово считалось правильно понятым аудитором, если он его написал правильно. Фраза считалась правильно понятой аудитором, если он правильно записал ее смысл (орфографические ошибки при этом не учитываются).

По результатам анализа распознанных аудиторами слов и фраз для каждого вида помехового сигнала и для каждого отношения с/ш рассчитывалась словесная ($W_{сл}$) и фразовая ($W_{фр}$) разборчивости речи:

$$W_{сл} = \frac{\sum_{i=1}^{k} N_i}{k \cdot N_{сл}} \quad (7)$$

$$W_{фр} = \frac{\sum_{i=1}^{k} M_i}{k \cdot M_{фр}} \quad (8)$$

где N_i – количество правильно распознанных слов i-аудитором;

M_i – количество правильно распознанных фраз i-аудитором;

$N_{сл}$ – общее количество слов, прослушиваемых аудиторами;

$M_{фр}$ – общее количество фраз, прослушиваемых аудиторами;

k – количество аудиторов.

Результаты расчетов представлены в таблице 3.

Зависимости словесной и фразовой разборчивости речи от отношения с/ш с высокой степенью достоверности аппроксимируются формулами:

$$W_{сл} \approx \Phi\left(Q_{сл.1} \cdot q - Q_{сл.2}\right) \quad (9)$$

$$W_{фр} \approx \Phi\left(Q_{фр.1} \cdot q - Q_{фр.2}\right) \quad (10)$$

где q – отношение сигнал/шум в частотном интервале 90–11200 Гц, дБ;

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$ – интеграл вероятности;

$\Phi^{-1}(x)$ – функция, обратная $\Phi(x)$;

$Q_{сл.1}$, $Q_{сл.2}$, $Q_{фр.1}$, $Q_{фр.2}$ – коэффициенты, характеризующие тестовые речевые сигналы, условия прослушивания и индивидуальные особенности аудиторов.

Выбор значений коэффициентов $Q_{сл.1}$, $Q_{сл.2}$, $Q_{фр.1}$, $Q_{фр.2}$ проводился из условия максимизации достоверности аппроксимации R^2, которая рассчитывалась по формуле [6]

$$R^2 = 1 - \frac{\sum_{i=1}^{m} \left(W_i - W_{э.i}\right)^2}{\sum_{i=1}^{m} \left(W_{э.i}\right)^2 - \frac{1}{m} \sum_{i=1}^{m} \left(W_{э.i}\right)^2} \quad (11)$$

где W_i – разборчивость речи для i-го отношения с/ш, рассчитанная по формуле (11) или (12);

$W_{э.i}$ – разборчивость речи для i-го отношения с/ш, рассчитанная по результатам экспериментальных исследований;

Таблица 4.

Значения коэффициентов $Q_{сл.1}$, $Q_{сл.2}$, $Q_{фр.1}$, $Q_{фр.2}$

Вид помехового сигнала	Словесная разборчивость речи			Фразовая разборчивость речи		
	$Q_{сл.1}$	$Q_{сл.2}$	R^2	$Q_{фр.1}$	$Q_{фр.2}$	R^2
«Белый» шум	0,13	1,45	1	0,21	2,98	1
«Розовый» шум	0,14	1,24	0,999	0,21	2,22	0.999
«Коричневый» шум	0,13	2,08	0,993	0,28	5,65	0.996
Шумовая «речеподобная» помеха	0,14	0,85	0,999	0,19	1,69	1

m – количество значений отношений с/ш, для которые разборчивость речи оценивалась экспериментально.

Рассчитанные значения коэффициентов $Q_{сл.1}$, $Q_{сл.2}$, $Q_{фр.1}$, $Q_{фр.2}$ приведены в таблице 4.

Графики зависимости разборчивости речи от отношения сигнал/шум для различного вида помех приведены на рисунках 1 и 2.

Точками (•) на рисунках обозначены значения разборчивости речи, полученные по результатам экспериментальных исследований.

Анализ рисунков показал, что разборчивость речи зависит не только от отношения сигнал/шум, но и от вида шума.

Для сравнительной оценки эффективности шумовых помех введем понятие «коэффициент эффективности шумовой помехи», под которым будем понимать отношение сигнал/шум для определенного вида шума к отношению сигнал/шум для «белого» шума, выраженных в дБ, при которой разборчивости речи одинаковы:

$$\kappa_{эф.ш} = \frac{q_{ш}}{q_{бш}} \tag{12}$$

$$\Phi^{-1}(W_{ш}) = \Phi^{-1}(W_{бш}) = Q_{1ш} \cdot (q_{ш} \cdot \kappa_{эф.ш}) - Q_{2ш} = Q_{1бш} \cdot q_{бш} - Q_{2бш} \tag{13}$$

где $\kappa_{эф.ш}$ – коэффициент эффективности шумовой помехи;

$q_{ш}$ – отношение сигнал/шум для определенного вида шума, при котором разборчивость речи равна $W_{ш}$, дБ;

$q_{бш}$ – отношение сигнал/шум для «белого» шума, при котором разборчивость речи равна $W_{бш} = W_{ш}$, дБ.

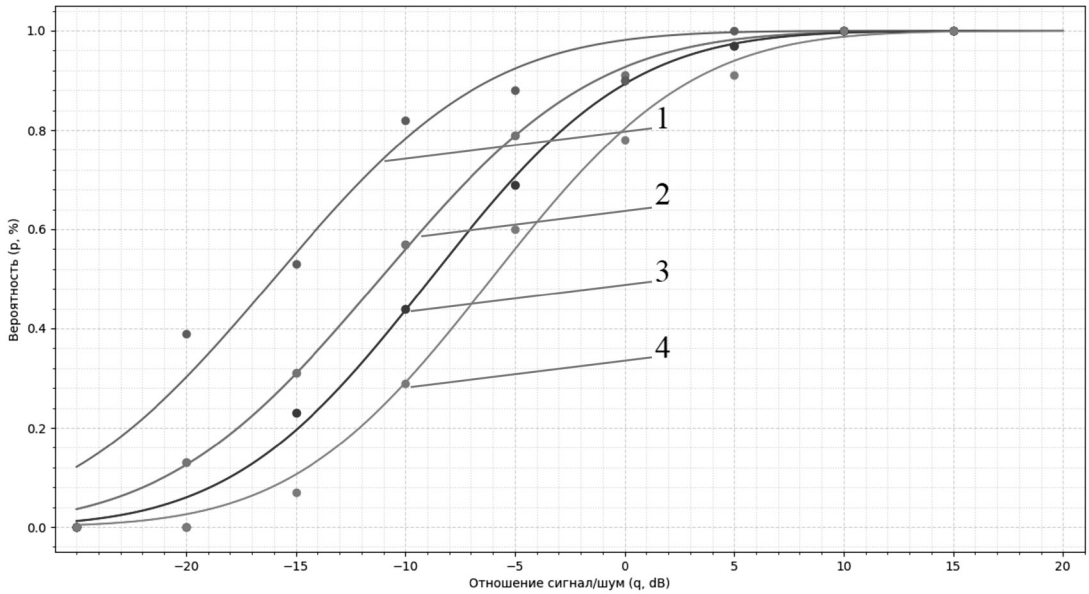

Рис. 1. Графики зависимости словесной разборчивости речи от отношения сигнал/шум:
1 – «коричневый» шум; 2 – «белый» шум; 3 – «розовый» шум; 4 – шумовая «речеподобная» помеха

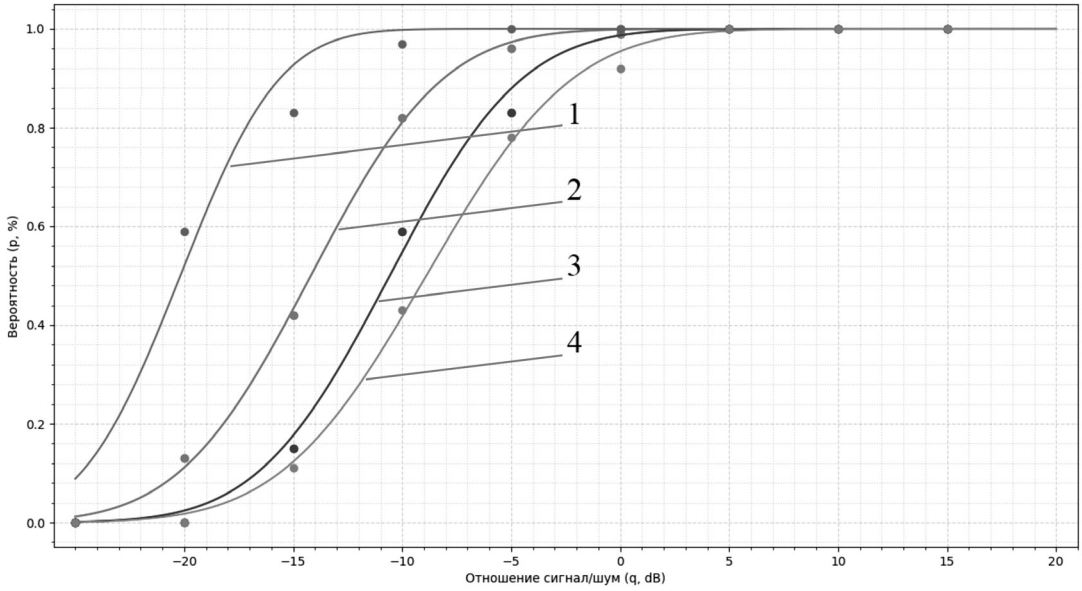

Рис. 2. Графики зависимости фразовой разборчивости речи от отношения сигнал/шум:
1 – «коричневый» шум; 2 – «белый» шум; 3 – «розовый» шум; 4 – шумовая «речеподобная» помеха

Из формулы (15) легко найди значение кэф.ш:

$$\kappa_{эф.ш} = \frac{Q_{1бш} \cdot q_{бш} + Q_{2ш} - Q_{2бш}}{Q_{1ш} \cdot q_{бш}} = \frac{Q_{1бш}}{Q_{1ш}} + \frac{Q_{2ш} - Q_{2бш}}{Q_{1ш} \cdot q_{бш}} \quad (14)$$

Рассчитанные значения коэффициентов $\kappa_{эф.ш}$ для различных видов шумов приведены в таблице 5.

Таблица 5.

Значения коэффициентов $\kappa_{эф.ш}$ для различных видов шумов

Вид помехового сигнала	Формулы для расчета коэффициента кэф.ш	
	Словесная разборчивость речи	Фразовая разборчивость речи
«Белый» шум	$\kappa_{эф.бш} = 1$	$\kappa_{эф.бш} = 1$
«Розовый» шум	$\kappa_{эф.рш} = 0{,}93 + 1{,}5/q$	$\kappa_{эф.рш} = 1{+}3{,}6/q$
«Коричневый» шум	$\kappa_{эф.кш} = 1 + 4{,}85/q$	$\kappa_{эф.кш} = 0{,}75{+}9{,}54/q$
Шумовая «речеподобная» помеха	$\kappa_{эф.рп} = 0{,}93 - 4{,}29/q$	$\kappa_{эф.рп} = 1{,}1 - 6{,}79/q$

Графические зависимости коэффициентов кэф.ш для различных видов шумов приведены на рисунках 3 и 4.

Таким образом формулы для расчета словесной и фразовой разборчивости речи для различных видов шумов будут иметь вид:

$$W_{сл.ш} \approx \Phi\left(Q_{сл.1.ш} \cdot \kappa_{эф.ш} \cdot q - Q_{сл.ш.2}\right) \quad (15)$$

$$W_{фр.ш} \approx \Phi\left(Q_{фр.1.ш} \cdot \kappa_{эф.ш} \cdot q - Q_{фр.ш.2}\right) \quad (16)$$

где кэф.ш – коэффициент эффективности шумовой помехи, значение которого приведено в таблице 5;

$Q_{сл.1}$, $Q_{сл.2}$, $Q_{фр.1}$, $Q_{фр.2}$ – коэффициенты, значения которых приведены в таблице 3.

3. Методика обоснования критериев эффективности защиты речевой информации от ее утечки по техническим каналам

Так как длительность мероприятий конфиденциального характера может изменяться в широких пределах, целесообразно перехваченный разговор разбить на фрагменты, например, длительностью до 10 мин.

Учитывая, что средний темп речи на русском языке 120 –130 слов в минуту [4], за 10 минут можно в среднем произнести 1200 – 1300 слов или 140 – 180 коротких фраз.

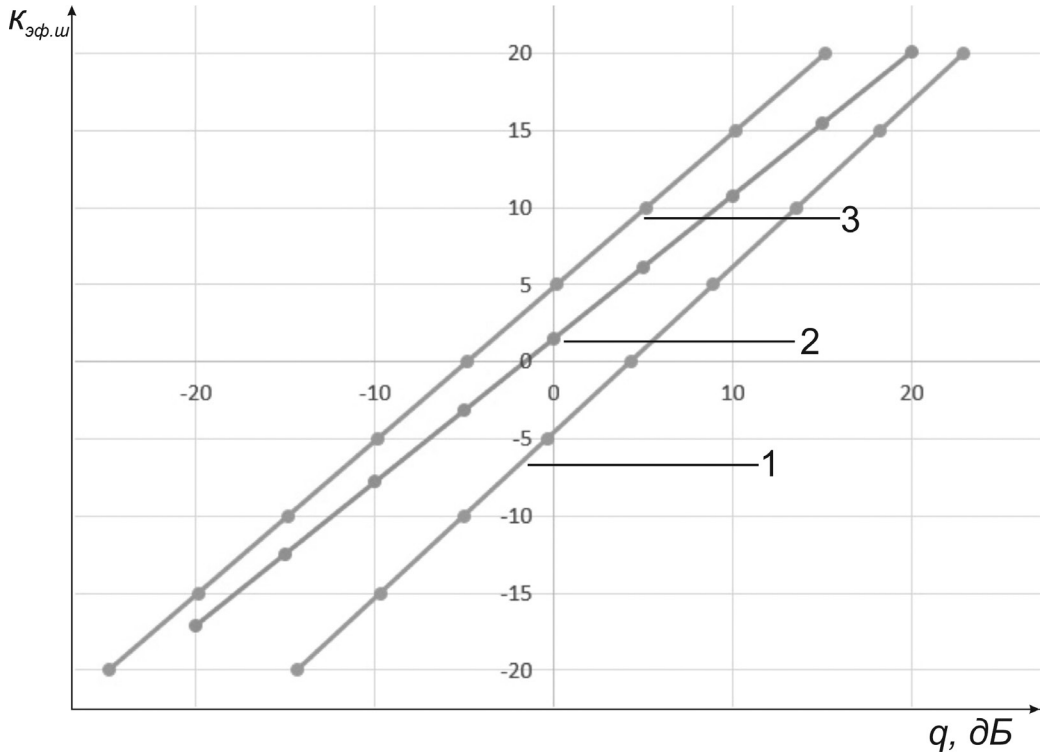

Рис. 3. Зависимости коэффициента эффективности шумовой помехи
для словесной разборчивостей речи от вида шума:
1 – шумовая «речеподобная» помеха; 2 – «розовый» шум; 3 – «коричневый» шум

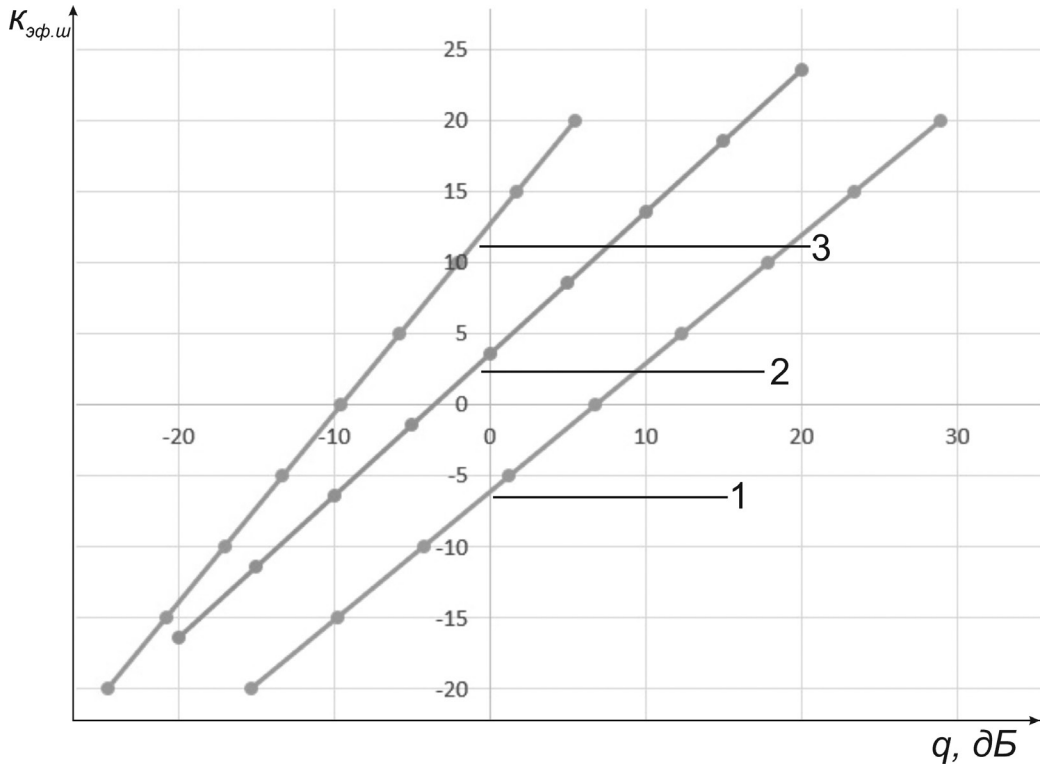

Рис. 4. Зависимости коэффициента эффективности шумовой помехи
для фразовой разборчивостей речи от вида шума:
1 – шумовая «речеподобная» помеха; 2 – «розовый» шум; 3 – «коричневый» шум

Для каждого фрагмента разговора необходимо определить его тематику и составить аннотацию.

Будем полагать, что цели защиты речевой информации достигнуты, если для каждого фрагмента разговора нельзя соответственно определить его тематику, или нельзя составить его аннотацию.

С учетом вышесказанного, авторами предлагается следующая методика обоснования критериев эффективности защиты речевой информации от ее утечки по техническим каналам:

1) Экспериментально определяется относительное среднее количество ключевых слов $k_{сл}$ и ключевых фраз $k_{фр}$ по различной тематике.

2) По каждой тематике экспериментально определяются среднее количество ключевых слов $N_{сл.кл}$ и ключевых фраз $N_{фр.кл}$ в разговоре, длительностью 10 минут.

3) По каждой тематике экспериментально определяются минимальное количество ключевых слов $N_{сл.кл.min}$, необходимых для вскрытия тематики разговора длительностью 10 минут, и минимальное количество ключевых фраз $N_{фр.кл.min}$, необходимых для составления его аннотации.

4) Экспертным методом определяются пороговые значения вероятности вскрытия тематики перехваченного разговора $P_{т.п}$, длительностью 10 минут, и вероятности составления аннотации его аннотации $P_{ан.п.}$

5) По формулам (1) и (2) строится зависимость вероятности вскрытия тематики перехваченного разговора $P_т$, длительностью 10 минут, от словесной разборчивости речи $W_{сл.}$.

6) По графику зависимости вероятности вскрытия тематики перехваченного разговора $P_т$ от словесной разборчивости речи $W_{сл}$ для порогового значения вероятности вскрытия тематики перехваченного разговора $P_{т.п}$ определяется пороговое значение словесной разборчивости речи $W_{сл.п.}$

7) По формулам (5) и (6) строится зависимость вероятности составления аннотации

Таблица 6.

Зависимости вероятностей скрытия тематики разговора ($P_т$) и составления его аннотации ($Ран$) от словесной ($W_{сл}$) и фразовой ($W_{фр}$) разборчивости речи от распознанных ключевых слов ($N_{сл.кл.р}$) и ключевых фраз ($N_{фр.кл.р}$)

$W_{сл}$	Вероятность вскрытия тематики разговора $P_т$				$W_{фр}$	Вероятность составления аннотации разговора $P_{ан}$			
	$N_{сл.кл.р} \geq 4$	$N_{сл.кл.р} \geq 5$	$N_{сл.кл.р} \geq 6$	$N_{сл.кл.р} \geq 7$		$N_{фр.кл.р} \geq 7$	$N_{фр.кл.р} \geq 8$	$N_{фр.кл.р} \geq 9$	$N_{фр.кл.р} \geq 10$
0,1	0,4	0,209	0,092	0,034	0,1	0	0	0	0
0,15	0,731	0,537	0,345	0,192	0,15	0,003	0	0	0
0,2	0,91	0,799	0,642	0,465	0,2	0,02	0,004	0,001	0
0,25	0,976	0,933	0,85	0,725	0,25	0,069	0,021	0,005	0,001
0,3	0,995	0,982	0,951	0,89	0,3	0,164	0,064	0,02	0,005
0,35	0,999	0,996	0,987	0,965	0,35	0,304	0,147	0,057	0,017
0,4	1	0,999	0,997	0,992	0,4	0,473	0,274	0,13	0,049
0,45	1	1	1	0,998	0,45	0,642	0,434	0,245	0,112
0,5	1	1	1	1	0,5	0,785	0,603	0,397	0,215
0,55	1	1	1	1	0,55	0,888	0,755	0,566	0,358
0,6	1	1	1	1	0,6	0,951	0,87	0,726	0,527
0,65	1	1	1	1	0,65	0,983	0,943	0,853	0,696
0,7	1	1	1	1	0,7	0,995	0,98	0,936	0,836
0,75	1	1	1	1	0,75	0,999	0,995	0,979	0,931
0,8	1	1	1	1	0,8	1	0,999	0,996	0,98
0,85	1	1	1	1	0,85	1	1	1	0,997
0,9	1	1	1	1	0,9	1	1	1	1
0,95	1	1	1	1	0,95	1	1	1	1
1	1	1	1	1	1	1	1	1	1

Таблица 7.

Показатели и критерии эффективности защиты речевой информации от ее утечки по техническим каналам

Цели защиты речевой информации	Показатели эффективности защиты речевой информации	Критерии эффективности защиты речевой информации		
		Пороговая вероятность скрытия тематики разговора $P_{т.п}$	Пороговая словесная разборчивость речи $W_{сл.п}$	Пороговая фразовая разборчивость речи $W_{фр.п}$
Скрытие тематики текста	Вероятность скрытия тематики разговора $P_т$	0,1	0,12	-
		0,2	0,15	-
		0,3	0,17	-
		0,4	0,2	-
Скрытие содержания текста	Вероятность составления аннотации разговора $P_{ан}$	0,1	-	0,45
		0,2	-	0,5
		0,3	-	0,53
		0,4	-	0,56

разговора $P_{ан}$, длительностью 10 минут, от фразовой разборчивости речи $W_{фр}$.

8) По графику зависимости вероятности составления аннотации разговора $P_{ан}$, длительностью 10 минут, от фразовой разборчивости речи $W_{фр}$ для порогового значения вероятности составления аннотации разговора Ран.п определяется пороговое значение фразовой разборчивости речи $W_{фр.п.}$

Таким образом, для обоснования критериев эффективности защиты речевой информации необходимо рассчитать словесную и фразовую разборчивости речи и оценить среднее количество слов и фраз в перехваченном разговоре.

Среднее количество ключевых слов и ключевых фраз зависит от тематики разговора. Например, по результатам анализа научных докладов по тематике «информационная безопасность» среднее значение ключевых слов составляло 2,4%, по данным, проведенным в работе [5], и 1,97% – по данным приведенным в работе [6]. При этом среднее значение ключевых фраз составило 10,3% по данным, проведенным в работе [5], и 9,67% – по данным приведенным в работе [6].

В качестве примера в таблице 6 приведены зависимости вероятностей вскрытия тематики разговора ($P_т$) и составления его аннотации ($P_{ан}$) от словесной ($W_{сл}$) и фразовой ($W_{фр}$) разборчивости речи от распознанных ключевых слов ($N_{сл.кл.р}$) и ключевых фраз ($N_{фр.кл.р}$) для разговора, длительностью 10 минут, состоящего из 1300 слов (2,5% из которых ключевые) и 160 фраз (10% из которых ключевые).

Задаваясь пороговыми значениями вероятности скрытия тематики разговора $P_{т.п}$ составления его аннотации, а также значениями минимального количества ключевых слов $N_{сл.кл.min}$ и ключевых фраз $N_{фр.кл.min}$ по таблице легко определить пороговые значения словесной $W_{сл.п}$ и фразовой $W_{фр.п}$ разборчивости речи.

В таблице 7 приведены пороговые значения словесной и фразовой разборчивости речи для достижения требуемого уровня защиты информации для разговора, длительностью 10 минут, состоящего из 1300 слов (2,5% из которых ключевые) и 160 фраз (10% из которых ключевые) при $N_{сл.кл.min}$ = 7, а $N_{фр.кл.min}$ = 10.

Анализ результатов экспериментальных исследований, представленные в п. 2, показал, что фразовая разборчивость речи ($W_{фр}$) связана со словесной разборчивостью ($W_{сл}$) зависимостью, представленной на рисунке 6, которую можно аппроксимировать функцией:

$$W_{фр} \approx \Phi\left(5{,}15 \cdot W_{сл} - 1{,}45\right) \qquad (17)$$

где $\Phi(x) = \frac{1}{2\pi} \cdot \int_{-\infty}^{x} exp\left(-\frac{t^2}{2}\right) dt$ – интеграл вероятности.

Представленная на рисунке 5 зависимость фразовой разборчивости речи от словесной согласуется с результатами экспериментальных исследований профессора Покровского Н.Б. (на графике 5 точками (•) представлены результаты экспериментальных исследований профессора Покровского Н.Б. [7]).

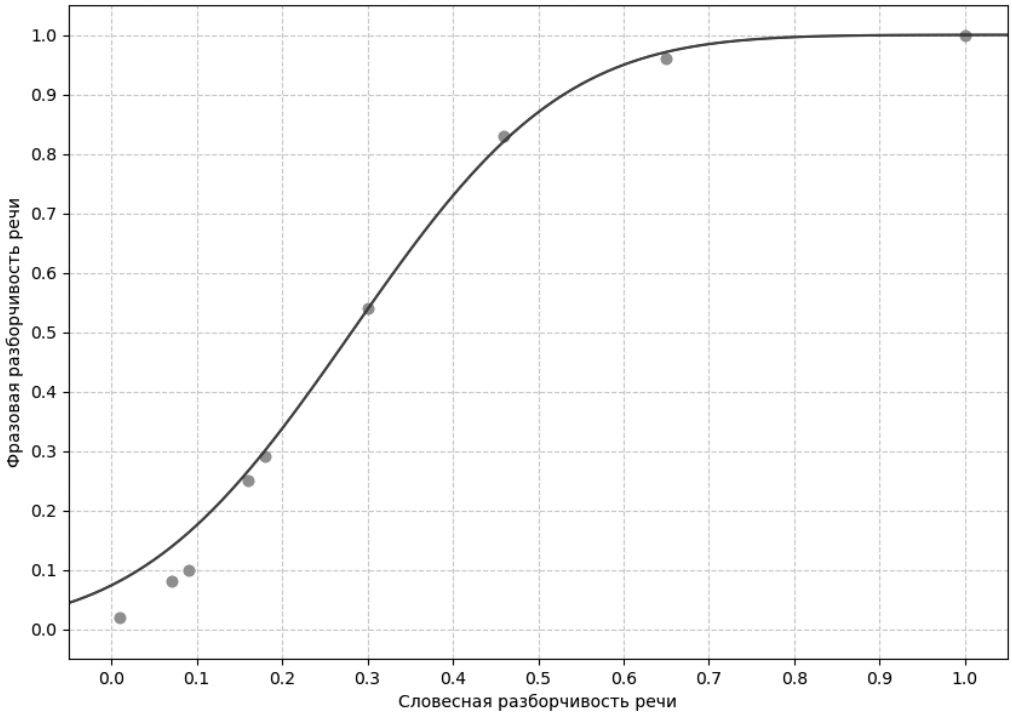

Рис. 5. Зависимость фразовой разборчивости речи ($W_{фр}$) от словесной ($W_{сл}$), рассчитанная по формуле (17)

Из формулы (17) легко найти значение словесной разборчивости речи $W_{сл}$:

$$W_{сл} \approx 0{,}194 \cdot \left[\Phi^{-1}\left(W_{фр}\right) + 1{,}45 \right] \qquad (18)$$

где $\Phi^{-1}(x)$ – функция, обратная интегралу вероятности.

С учетом формулы (18) можно получить зависимость вероятности составления аннотации текста от словесной разборчивости речи для различного минимального количества ключевых фраз ($N_{фр.кл.min}$).

В табл. 8 приведены показатели и критерии эффективности защиты речевой информации от ее утечки по техническим каналам с учетом пересчета фразовой разборчивости текста в словесную.

Анализ данных, представленных в таблице 8, показывает их корреляцию с данными, представленными в таблице 1.

Полученные результаты исследований могут быть использованы при обосновании показателей и критериев эффективности защиты речевой информации в выделенных помещениях.

Таблица 8.

Показатели и критерии эффективности защиты речевой информации от ее утечки по техническим каналам

Цели защиты речевой информации	Показатели эффективности защиты речевой информации	Критерии эффективности защиты речевой информации	
		Пороговая вероятность скрытия тематики разговора $P_{т.п}$	Пороговая словесная разборчивость речи $W_{сл.п}$
Скрытие тематики текста	Вероятность скрытия тематики разговора $P_т$	0,1	0,12
		0,2	0,15
		0,3	0,17
		0,4	0,2
Скрытие содержания текста	Вероятность составления аннотации разговора $P_{ан}$	0,1	0,26
		0,2	0,28
		0,3	0,3
		0,4	0,31

Литература

1. Дворянкин С.В., Макаров Ю.К., Хорев А.А. Обоснование критериев эффективности защиты речевой информации// Защита информации. Инсайд. – С. Петербург.: 2007. – № 2 – С. 18 – 25.

2. Вентцель Е. С. Задачи и упражнения по теории вероятностей: Учеб. пособие для студ. втузов / Е. С. Вентцель, Л. А. Овчаров. - 5-е изд., испр. - М.: Издательский центр «Академия», 2003. - 448 с.

3. Яремко Н. Н. Краткий курс комбинаторики, теории вероятностей и математической статистики: учеб. пособие / Н. Н. Яремко, О. Г. Никитина. – Пенза: Изд-во ПГУ, 2017. – 134 с.

4. Метлова В. А. Темп речи в свободной коммуникации: социолингвистический аспект// Вестник Пермского университета Российская и зарубежная филология. Пермь: Изд-во Пермск. ун-та, 2014. Вып. – 4 (28). – С. 58-65.

5. Хорев А.А., Суворова В.Ю. Оценка относительного количества ключевых фраз и слов в научных текстах// Всероссийская конференция «Радиоэлектронные устройства и системы для инфотелекоммуникационных технологий – РЭУС-2023». Доклады. – М.: РНТОРЭС имени А.С.Попова. 2023. – С. 412 – 415.

6. Хабирова Л.А. Оценка среднего количества ключевых слов и ключевых фраз в научных текстах: выпускная квалификационная работа: 10.03.01; рук. Хорев А.А. – М.: МИЭТ, 2024. – 72 с.

7. Покровский Н.Б. Расчет и измерение разборчивости речи. – М.: Гос. Издательство литературы по вопросам связи и радио, 1962. – 392 с.

References

1. Dvoryankin S.V., Makarov YU.K., Khorev A.A. Obosnovaniye kriteriyev effektivnosti zashchity rechevoy informatsii// Zashchita informatsii. Insayd. – S. Peterburg.: 2007. – № 2 – S. 18 – 25.

2. Venttsel' Ye. S. Zadachi i uprazhneniya po teorii veroyatnostey: Ucheb. posobiye dlya stud. vtuzov / Ye. S. Venttsel', L. A. Ovcharov. - 5-ye izd., ispr. - M.: Izdatel'skiy tsentr «Akademiya», 2003. - 448 s.

3. Yaremko N. N. Kratkiy kurs kombinatoriki, teorii veroyatnostey i matematicheskoy statistiki: ucheb. posobiye / N. N. Yaremko, O. G. Nikitina. – Penza: Izd-vo PGU, 2017. – 134 s.

4. Metlova V. A. Temp rechi v svobodnoy kommunikatsii: sotsiolingvisticheskiy aspekt// Vestnik Permskogo universiteta Rossiyskaya i zarubezhnaya filologiya. Perm': Izd-vo Permsk. un-ta, 2014. Vyp. – 4 (28). – S. 58-65.

5. Khorev A.A., Suvorova V.YU. Otsenka otnositel'nogo kolichestva klyuchevykh fraz i slov v nauchnykh tekstakh// Vserossiyskaya konferentsiya «Radioelektronnyye ustroystva i sistemy dlya infotelekommunikatsionnykh tekhnologiy – REUS-2023». Doklady. – M.: RNTORES imeni A.S.Popova. 2023. – S. 412 – 415.

6. Khabirova L.A. Otsenka srednego kolichestva klyuchevykh slov i klyuchevykh fraz v nauchnykh tekstakh: vypusknaya kvalifikatsionnaya rabota: 10.03.01; ruk. Khorev A.A. – M.: MIET, 2024. – 72 s.

7. Pokrovskiy N.B. Raschet i izmereniye razborchivosti rechi. – M.: Gos. Izdatel'stvo literatury po voprosam svyazi i radio, 1962. – 392 s.

ХОРЕВ Анатолий Анатольевич, доктор технических наук, профессор, заведующий кафедрой «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Московский институт электронной техники» (МИЭТ). 124498, Москва, г. Зеленоград, площадь Шокина, дом 1. E-mail: horev@miee.ru

ПОРСЕВ Илья Сергеевич, старший преподаватель кафедры «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Московский институт электронной техники» (МИЭТ). 124498, Москва, г. Зеленоград, площадь Шокина, дом 1. E-mail: ib.labs@yandex.ru

HOREV Anatoly Anatolyevich, Doctor of Technical Sciences, Professor, Head of the Department of Information Security, Federal State Autonomous Educational Institution of Higher Education, National Research University Moscow Institute of Electronic Technology (MIET). 124498, Moscow, Zelenograd, Shokin Square, Building 1. E-mail: horev@miee.ru

PORSEV Ilya Sergeevich, Senior Lecturer, Department of Information Security, Federal State Autonomous Educational Institution of Higher Education, National Research University Moscow Institute of Electronic Technology (MIET). 124498, Moscow, Zelenograd, Shokin Square, Building 1. E-mail: ib.labs@yandex.ru

УДК 004.056 Вестник УрФО № 2(52) / 2024, с. 37–45

Котельников Н. Д., Афанасьева М. В., Баранкова И. И.

DOI: 10.14529/secur240204

ПРИМЕНЕНИЕ ДВУСТОРОННЕЙ СИГНАЛЬНОЙ ИГРЫ В ТЕХНОЛОГИЯХ DECEPTION ДЛЯ ВЫБОРА ОПТИМАЛЬНОЙ СТРАТЕГИИ ЗАЩИТЫ

Все чаще для борьбы с сетевыми атаками сотрудники кибербезопасности прибегают к использованию методов киберобмана (deception). Технология deception, использующая ложные данные для привлечения злоумышленников, позволяет выявлять попытки нежелательного доступа к сети и предотвратить атаку на ранних стадиях.

Несмотря на эффективность средств киберобмана, их развертывание является сложным и дорогостоящим. Необходимо понимать, при каких условиях использование данных технологий будет наиболее целесообразным.

Использование теории игр помогает анализировать стратегии злоумышленников и разрабатывать наиболее эффективные методы защиты. Одним из способов определения оптимальной стратегии защиты является динамическая игра с неполной информацией (сигнальная игра), где один из игроков (атакующий или защитник) является отправителем сигнала, а второй игрок – получателем.

В статье предлагается расширить сигнальную игру до двусторонней: рассмотреть действия, как злоумышленников, так и сотрудников кибербезопасности, рассчитать возможные их выгоды от осуществления атаки и реализации системы защиты. Данная модель позволит оценить необходимость использования средств киберобмана и выбрать наиболее эффективную стратегию защиты.

Ключевые слова: *информационная безопасность, сетевая атака, технология обмана, стратегия защиты, теория игр, сигнальная игра.*

Kotelnikov N. D., Afanasyeva M. V., Barankova I. I.

APPLICATION OF TWO-SIDED SIGNALING GAME IN DECEPTION TECHNOLOGIES TO SELECT THE OPTIMAL DEFENSE STRATEGY

Increasingly, cybersecurity officials are turning to deception techniques to combat network attacks. Deception, which uses false data to lure attackers, can detect unwanted network access attempts and prevent attacks early on.

Despite the effectiveness of cyber deception tools, their deploying is complex and costly. It is important to understand under what conditions the use of these technologies will make the most sense.

The use of game theory helps analyze attacker strategies and develop the most effective defense techniques. One way to determine the optimal defense strategy is a dynamic game with incomplete information (signaling game), where one of the players (attacker or defender) is the sender of the signal, and the second player is the receiver.

The paper proposes to extend the signaling game to a two-sided game: consider the actions of both attackers and cybersecurity personnel, and calculate their possible benefits from the attack and the implementation of the defense systemThis model will allow us to assess the need to use cyber defenses and choose the most effective defense strategy.

Keywords: *information security, network attack, deception technology, defence strategy, game theory, signal game.*

Введение

Рост киберпреступности наносит огромный денежный и репутационный ущерб предприятия, нивелируя выгоду от использования информационных систем.

В целях защиты от компьютерных вторжений и обнаружения кибератак на ранних этапах была разработана технология deception, позволяющая запутать или ввести в заблуждение злоумышленников. В основе deception лежат техники обмана злоумышленников путем применения ловушек, приманок и других способов запутывания преступников [1]. Цель защитников заключается в предоставлении ложной информации, способной вывести атакующего «из игры». Злоумышленникам приходится тратить дополнительные ресурсы (например, время и деньги). Защитники же получают оперативную информацию о проникновении и реагируют на любые действия атакующего[2].

Развертывание средств обмана является дорогостоящей и сложной задачей: злоумышленник может найти и обойти ловушку или использовать её против защищающегося. Необходимо, чтобы ловушка обеспечивала максимальную вероятность того, что злоумышленник совершит атаку на неё, а не на реальную систему. Необходимо выбрать правильную последовательность защитных мер, предсказав действия атакующего, предотвратить процесс атаки и максимизировать свою выгоду. Методы теории игр помогают проверить и повысить эффективность технологий обмана. В [3] затрагивается вопрос выбора стратегии ложной информационный системы на основе модели теории игр. В работе учитываются показатели игроков и все возможные ограничения.

В [4] разработана биматричная игровая модель взаимодействия информационной системы с агрессивной внешней средой. Отдельно в работе рассматриваются возможные варианты трансформации базовой статической биматричной модели в динамическую игру с неполной информацией.

Применение сигнальных игр в deception технологиях обсуждалось в [5]. В работе [6] сигнальная игра была рассмотрена со стороны атакующего и защитника, с целью определения возможных действий нарушителя.

В статье предложен метод определения оптимальной стратегии, основанный на двусторонней сигнальной игре, который позволит эффективно применять технологии киберобмана в качестве защиты сети. Существенным отличием от методов [5,6] является то, что игра, предложенная в статье, является многоэтапной и двусторонней, где атакующий и защитник, наблюдая за действиями конкурента, постоянно изменяют стратегию своих действий.

Сигнальная игра

Сигнальная игра характеризуются неполнотой информации о соперниках. Злоумышленник не может понять, насколько защищенной является система, которую он собирается атаковать. Защитник не имеет представления о нападающем, о его возможностях и о технологиях, доступных ему.

Иными словами, в задачах подобного рода выбор решения зависит от состояний объективной действительности, называемой в модели «природой». [7]. Природа не является разумным игроком, она не заинтересована в выигрыше. Её роль – задать начальные условия, по которым будет проходить игра в дальнейшем. В сигнальной игре Природа наделяет отправителя сигнала типом θ. Знание о типе соперника дает информацию о его стратегиях и выигрышах, позволяет выбрать оптимальное решение. Однако Получатель сигнала не знает о типе отправителя, а лишь имеет априорное предположение $P(\theta)$, вероятность того, что отправитель принадлежит к данному типу.

В реальности игра носит двусторонний характер, где атакующий и защитник, наблюдая за действиями конкурента, изменяют стратегию своих действий. На рисунке 1 изображен данный процесс.

На этапе 1 защитник выступает в роли отправителя сигнала, а атакующий в роли получателя. Защитник развертывает сеть, настраивает топологию, задает IP-адреса. Злоумышленник различными способами пытается собрать информацию о сети. Такая информация служит основой для проведения сетевой атаки. В нашем случае информация рассматривается как сигнал M_d, выпущенный защитником. Атакующий наблюдает за сигналом M_d, и выбирает наиболее выигрышное ответное действие (оптимальную стратегию атаки).

По ходу игры стороны нападения и защиты постоянно меняются ролями. Каждый этап игры состоит из базовой игры с сигналами, как показано на этапах 2, 3 и i на рисунке 1. На этапе 2 атакующий выбирает стратегию атаки и подает сигнал M_a. Защитник получает сигнал M_a, корректирует априорное суждение о типе атакующего и выбирает соответствующую стратегию защиты. В процессе противостояния сигнал передается в обоих направлениях, и стороны нападения и защиты используют правило Байеса для постепенной корректировки своих оценок истинного типа другой стороны. С точки зрения защитника, условием завершения игры является прекращение атаки со стороны нападающего и прекращение передачи сигналов (этап n на рисунке 1).

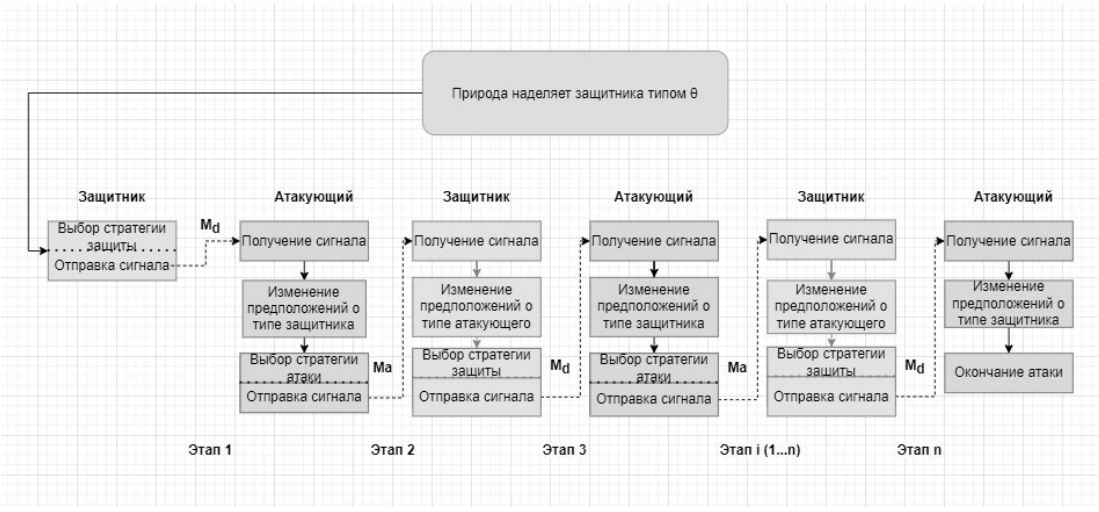

Рис. 1. Многоэтапная двусторонняя сигнальная игра

Построение двусторонней сигнальной модели

Двусторонняя сигнальная модель – это конечная игра, состоящая из нескольких сигнальных игр. В ДСМ атакующий и защитник попеременно выступают в роли отправителя и получателя сигналов.

В двусторонней сигнальной модели (ДСМ) будет использоваться 10 переменных

1) Игроки $N = (N_a, N_d)$, где N_d – защитник, а N_a – атакующий

2) Типы игроков $T = (T_a, T_d)$. тип атакующего $T_a = (\eta_h$ – продвинутый, η_l - обыкновенный); тип защитника $T_d = (\theta_h$ – усиленная система защиты, θ_l – обыкновенная система защиты);

3) $M = (M_a, M_d)$ – сигнал, который подают защитник и атакующий. Сигнал атакующего $M_a = (m_{ah}$ – продвинутый сигнал , m_{al} – обыкновенный сигнал); сигнал защитника M_d = (mdh – усиленный защитный сигнал, m_{dl} – обыкновенный защитный сигнал). Стоит отметить, что в ходе игры отправитель сигнала может попытаться обмануть оппонента, подавая сигналы, которые не соответствуют его собственному типу, с целью ввести соперника в заблуждение, заставив принять неправильный выбор.

4) $S = (D,A)$ – стратегии защитника и атакующего. Стратегия защиты $D = \{d_g | g=1,2,3\ldots.\}$ и стратегия атаки $A = \{a_h | h=1,2,3\ldots.\}$

5) K – стадия игры. $K = (1,2,3\ldots k)$. Сигнальная игра ведется в несколько этапов, и этап игры k представлен в виде ДСМ(k).

6) δ – коэффициент ослабления обмана по ходу игры. После многократных стратегических взаимодействий между атакующим и защищающимся стороны начинают узнавать типы друг друга и сигнал обмана постепенно ослабевает.

7) $P = (P_a, P_d)$ априорная вероятность. P_a - набор убеждений атакующего относительно типа защищающегося. P_d - является набором убеждений защитника относительно типа атакующего.

8) $P = (P_a', P_d')$ апостериорная вероятность. P_a' - является апостериорным набором предположений атакующего о типе защитника. P_d' - является апостериорным набором предположений защитника о типе атакующего.

9) ξ – коэффициент дисконтирования, определяющий величину прибыли на k-этапе.

10) $U = (U_d, U_a)$ обозначает ожидаемый набор выгод для защитника и атакующего.

Определение байесовского равновесия

Для определения стратегии защиты, необходимо найти байесовское равновесие.

В ДСМ(k) равновесие $EQ_k = (m^*(\theta)\ f^*(m),\ P_F'(T|m))$, где - $m^*(\theta)$ стратегия отправителя, $f^*(m)$, - стратегия получателя. $P_F'(T|m)$ – это апостериорная вероятность, которая вычисляется получателем сигнала в соответствии с байесовским правилом на основе априорной вероятности $P(\theta)$, сигнала m и оптимальной стратегии отправителя сигнала $m^*(\theta)$.

Согласно теории игр, равновесие должно удовлетворять условиям (1), (2):

$$f^*(m) \in \underset{f \epsilon F}{arg\,max} \sum P_F'(T|m) U_F(m^*(\theta), f, \theta) \quad (1)$$

$$m^*(\theta) \in \underset{m \epsilon M}{arg\,max}\, U_S(m, f^*(m), \theta) \quad (2)$$

$f^*(m)$ – оптимальная стратегия получателя сигнала после определения апостериорной вероятности $P_F'(\theta|m)$, о типе отправителя сигнала.

$m^*(\theta)$ – оптимальная стратегия отправителя сигнала, после прогнозирования оптимального действия $f^*(m)$ получателя.

Этапы достижения идеального байесовского следующие

Шаг 1. Вычисление оптимальной стратегии $f^*(m)$ на основе сигнала, который получает принимающий.

Шаг 2. Отправитель выбирает оптимальную стратегию $m^*(\theta)$.

Шаг 3. Рассчитать Байесовское равновесие: $EQ_k = (m^*(\theta)\ f^*(m),\ P_F'(T|m))$.

1) На первом этапе игры ДСМ(1) отправителем сигнала является защитник, а получателем атакующий.

Природа выбирает тип защитника. Тип θ_h выбирается с априорной вероятностью p_1, а тип θ_n – с вероятностью $1\text{-}p_1$. Защитник подает сигналы m_{dh} и m_{dl}. Если атакующий получает сигнал защиты m_{dh}, апостериорное заключение атакующего о типе защитника равно $(\alpha_1, 1\text{-}\alpha_1)$, а при получении сигнала m_{dl}, апостериорное заключение атакующего о типе защитника равно $(\beta_1, 1\text{-}\beta_1)$. Затем злоумышленник выбирает стратегию η_h, η_l. Получаем равновесие $EQ1 = (m^*(\theta)\ a^*(m),\ P_a'(\theta))$ для ДСМ(1). На рисунке 2 изображено дерево для этапа ДСМ(1).

2) На втором этапе игры ДСМ(2) отправителем сигнала является атакующий, а получателем защитник. Атакующий выбирает стратегию атаки в соответствии с EQ_1 и посылает сигнал защитнику. В ходе игры ДСМ(1) сторо-

ны нападения и защиты достигли получили определенные данные друг о друге, а значит, сигнал обмана ослабевает. Выбор типа атакующего определяется коэффициентом ослабления сигнала обмана σ и вероятностью $EQ_1(P_F'(T))$ и записывается как $\delta EQ_1(P_F'(T))$. Атакующий выбирает η_h с вероятностью $\delta EQ_1(P_F'(T))$ и η_l с вероятностью $1\text{-}\delta EQ_1(P_F'(T))$. Дерево игры ДСМ(2) показано на рисунке 3.

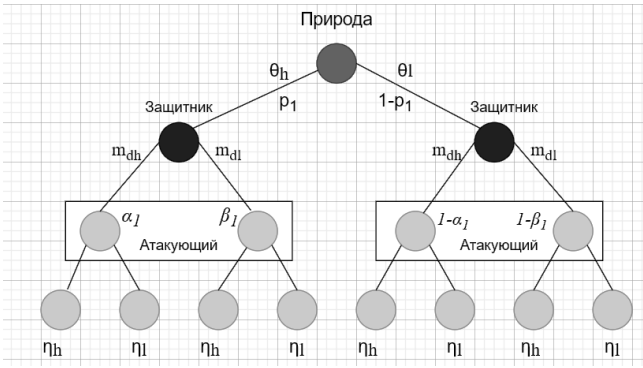

Рис. 2. Дерево для 1 этапа игры

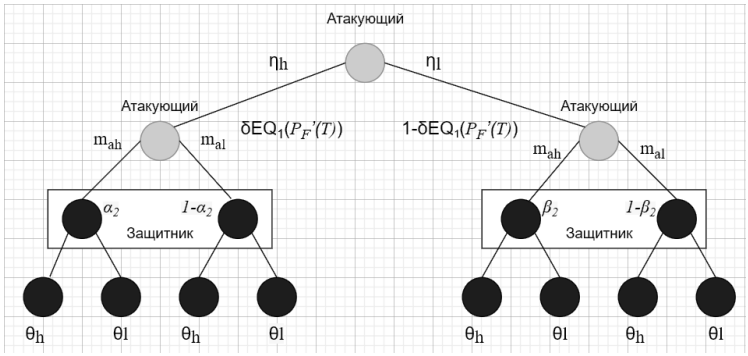

Рис. 3. Дерево для 2 этапа игры

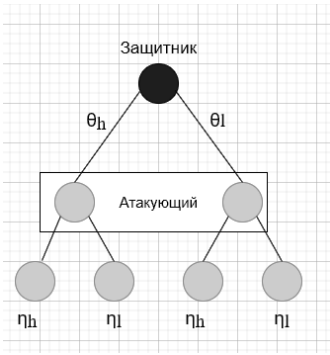

Рис. 4. Дерево для K этапа игры

Игроки на каждом этапе меняются ролями, а значение обманного сигнала уменьшается $\delta_k=\delta^{k-1}$.

На K-ом этапе игры, влияние обманного сигнала полностью исчезнет (например, при δ=0,4 на 7 этапе $\delta_8 = 0{,}4^6 = 0{,}004 \approx 0$: в данном случае K=7)

4) На K-этапе игры ДСМ(K) отправителем сигнала является защитник, а получателем атакующий. Дерево игры показано на рисунке 4.

На данном этапе, обман становится невозможен, поэтому подача ложных сигналов не имеет смысла.

Анализ результатов

Злоумышленник выбирает стратегию атаки, состоящей из нескольких последовательных действий (kill chain). Защитник же выбирает стратегию защиты. В таблице 1 показаны возможные стратегии атакующего и защитника и их расходы на реализацию этих стратегий.

В работе [6] представлены статистические данные для получения значения ущерба системе ($C_{ущ}$) при различных комбинациях стратегий атак и защиты (таблица 2).

Согласно методу расчета затрат [6], прибыль атакующего – это $C_{ущ}$, а затраты – сумма C_a и C_{ao}. Затраты защитника – это сумма $C_{ущ}$, $C_з$ и $C_{зо}$.

Функции прибыли атакующей U_a и обороняющейся U_d сторон могут быть выражены, соответственно, следующим образом (3,4):

$$U_a(a_h, d_g, k) = \sum_{g,h,k} \xi^{k-1} \left[C_{ущ} - C_a - C_{ao} + B_a\right] \quad (3)$$

$$U_d(a_h, d_g, k) = -\sum_{g,h,k} \xi^{k-1} \left[C_{ущ} + C_з + C_{зо} - B_з\right] \quad (4)$$

$B_з$ – выгода защитника при обмане атакующего (при подаче ложного сигнала). Поскольку злоумышленник тратит лишние ресурсы, у защитника появляется время на принятие решений и более эффективное противодействие атаке. $B_з = 100$

B_a – выгода атакующего при обмане защитника. $B_a = 50$

ξ – коэффициент дисконтирования равен 0,5.

δ – коэффициент ослабления сигнала равен 0,6.

Используем вышеописанный метод для получения наступательных и оборонительных выигрышей при различных комбинациях типов стратегий.

Таблица 1.

Стратегии атакующего и защитника

Действия злоумышленника	η_h		η_l		Действия защитника	θ_h		θ_l	
	a_1	a_2	a_3	a_4		d_1	d_2	d_3	d_4
Сканирование сети	+	+	+	+	Honeypot	+	+	+	+
Установка ВПО	+				Honey file	+	+	+	
Переполнение буфера	+	+	+		Антивирусная защита	+	+	+	+
Внедрение кода	+	+			Honey account	+	+		
Повышение привелегий	+	+			Виртуализация сети	+			
Брутфорс			+	+	Honeynet	+	+	+	
DDoS-атака	+		+	+	Разграничение доступа	+	+	+	+
Взлом учетной записи	+		+	+	Многофакторная аутентификация	+			
НСД к ресурсам	+	+			MTD	+	+		
Стоимость атаки (C_a)	**192**	**184**	**96**	**92**	**Стоимость защиты ($C_з$)**	**272**	**256**	**176**	**164**
Стоимость обмана (C_{ao})	**32**	**28**	**12**	**8**	**Стоимость обмана ($C_{зо}$)**	**40**	**32**	**16**	**12**

Таблица 2.

Стоимость ущерба системе

	d_1	d_2	d_3	d_4
a_1	928	952	1056	1072
a_2	908	892	1008	1028
a_3	872	848	912	928
a_4	848	832	884	904

Согласно [5] Природа выбирает тип защитника с вероятностями (0,4, 0,6). Предположим, что защитник с типом θ_h подает сигнал m_{ah}, а злоумышленник выбирает стратегию η_h. В таком случае, мы получаем 4 комбинации стратегий (a_1, d_1), $(a1, d_2)$, (a_2, d_1), (a_2, d_2). Показатели стоимости затрат взяты из таблицы 2, стоимость ущерба берется из таблицы 3.

Найдем функции полезности для атакующего и защитника при выборе комбинации (a_1, d_1) по формулам 5 и 6. Поскольку в данном случае нет обманного сигнала, то выгоды атакующего и защитника равны 0.

$$U_a(a_1 d_1 1) = C_{ущ}(a_1, d_1) - C_a - C_{ao} = 928 - 192 - 0 = 736 \quad (5)$$

В данном случае атакующий не подает сигнала, поэтому $C_{ao} = 0$.

$$U_d(a_1 d_1 1) = -\left[C_{ущ}(a_1, d_1) + C_з + C_{зо}\right] = -[928 + 272 + 0] = -1200 \quad (6)$$

Поскольку защитник не подает ложного сигнала $C_{зо} = 0$.

Так как нам неизвестно, какую стратегию выберет атакующий и защитник, предполагаем, что выбор каждой стратегии равновероятен.

Полученные значения представлены в таблице 3.

Аналогичным образом, используем метод для других комбинаций типов, стратегий

Таблица 3.

Выигрыши защитника и нападающего

	(a_1,d_1)	(a_1,d_2)	(a_2,d_1)	(a_2,d_2)	Ср. знач
U_a	736	724	760	708	732
U_d	-1184	-1114	-1225	-1091	-1184

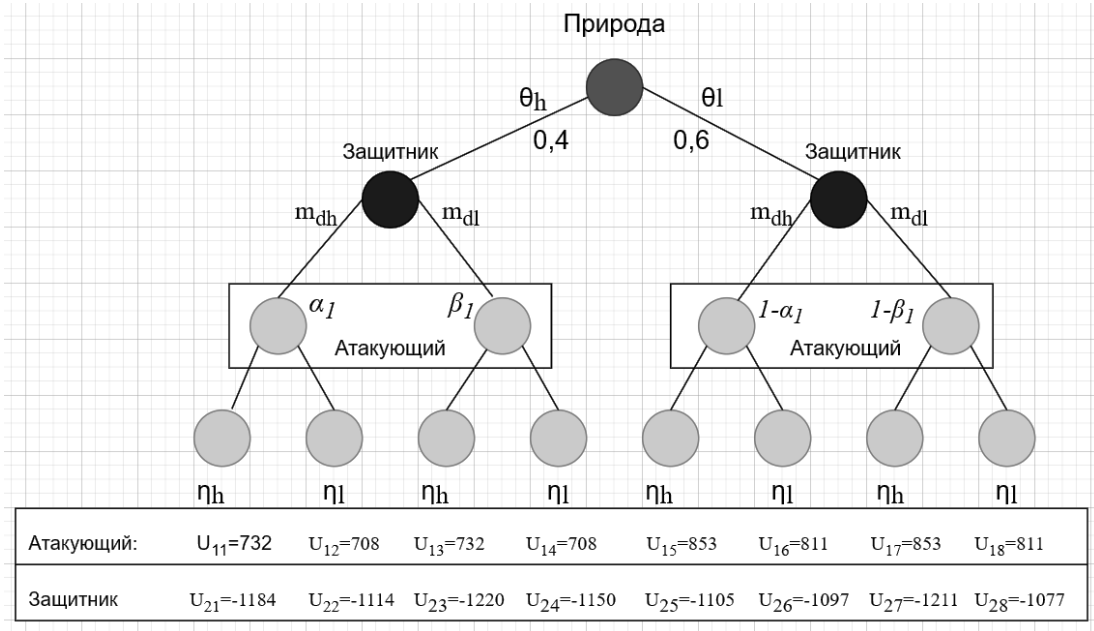

Рис. 5. Игровое дерево для 1 этапа игры ДСМ(1)

и сигналов. Результаты представлены на рисунке 5.

Используя алгоритм решения равновесия, для ДСМ(1) получаем две комбинации типов стратегий:

1) Когда защитник выбирает тип θ_h и подает сигнал m_{dh}, атакующий выбирает стратегию η_h. В этом случае $U_{a1} = -1184$ и $U_{d1} = 732$.

2) Защитник выбирает тип стратегии θ_l и подает сигнал m_{dh}. Па атакующий выбирает тип стратегии η_h. В таком случае $U_{a1} = -1105$ и $U_{d1} = 853$.

Защитник выбирает вариант 2 (θ_l, m_{dh}) в качестве стратегии защиты. Игровое дерево атаки и защиты показано на рисунке 5.

2) Результаты следующих этапов показаны в таблице 4.

Как видно из таблицы, на этапах со второго по шестой защитнику эффективнее использовать обманные сигналы, чтобы заставить злоумышленника выбрать неверную стратегию и уменьшить свои расходы на защиту. Однако с течением времени сигнал ослабевает. В этот момент обман не удастся, поскольку атакующий знает достаточно информации о защитнике. Поэтому, на данном этапе наиболее эффективной является применение усиленной системы защиты.

Предложенная модель двусторонней сигнальной игры имеет существенное значение в области информационной безопасности. С помощью предложенного метода защитник сети может выбрать наиболее эффективную защиту от сетевых атак.

Таблица 4.

Стратегия защиты на различных этапах

Этап	Роль защитника	Выигрыш защитника	Выигрыш атакующего	Стратегия защиты
1	Отправитель	-1105	853	(θ_l, m_{dh})
2	Получатель	-1114,2	821,4	θ_l
3	Отправитель	-1133,4	831,8	(θ_l, m_{dh})
4	Получатель	-1157,8	845	θ_l
5	Отправитель	-1168,2	858,2	(θ_l, m_{dh})
6	Получатель	-1182,6	827,8	θ_l
7	Отправитель	-984	804,4	(θ_h, m_{dh})

Литература

1. Исследование применения технологии deception для предотвращения угроз кибербезопасности. / Путято М. М., Макарян А. С., Чич Ш. М., Маркова В. К. // Прикаспийский журнал: управление и высокие технологии. 2020. №4. Стр. 85-98. URL: https://cyberleninka.ru/article/n/issledovanie-primeneniya-tehnologii-deception-dlya-predotvrascheniya-ugroz-kiberbezopasnosti (дата обращения: 03.04.2024).

2. Афанасьева М.В. Применение данных HONEYPOT-систем для прескриптивной аналитики действий злоумышленника. // Актуальные проблемы современной науки, техники и образования. 2023.№1. С.417

3. Шматова Е. С. Выбор стратегии ложной информационный системы на основе модели теории игр // Вопросы кибербезопасности. 2015. №5 (13). URL: https://cyberleninka.ru/article/n/vybor-strategii-lozhnoy-informatsionnyy-sistemy-na-osnove-modeli-teorii-igr (дата обращения: 13.04.2024).

4. Конюховский П.В. , Шабалин А.А. Теоретико-игровые подходы в анализе стратегий защиты корпоративных информационных систем // International Journal of Open Information Technologies. 2023 №12. URL: https://cyberleninka.ru/article/n/teoretiko-igrovye-podhody-v-analize-strategiy-zaschity-korporativnyh-informatsionnyh-sistem (дата обращения: 12.04.2024).

5. J. Pawlick and Q. Zhu, "Deception by design: evidence-based signaling games for network defense," 2015. URL: https://archive.org/details/arxiv-1503.05458/page/n23/mode/2up (дата обращения: 02.04.2024).

6. Y. Hu, H. Zhang, Y. Guo, T. Li, and J Ma, "A novel attack-and-defense signaling game for optimal deceptive defense strategy choice," Wireless Communications and Mobile Computing, vol. 2020. URL: https://www.hindawi.com/journals/wcmc/2020/8850356/ (date of application: 10.04.2024).

7. Голдуева Д.А., Мокшанина М. А. Принятие решений в условиях неопределенности на основе теории игр с природой. // Вестник ПензГУ. 2020. №4 (32). URL: https://cyberleninka.ru/article/n/prinyatie-resheniy-v-uslovii-neopredelennosti-na-osnove-teorii-igr-s-prirodoy (дата обращения: 14.04.2024).

References

1. Issledovaniye primeneniya tekhnologii deception dlya predotvrashcheniya ugroz kiberbezopasnosti. / Putyato M. M., Makaryan A. S., Chich SH. M., Markova V. K. // Prikaspiyskiy zhurnal: upravleniye i vysokiye tekhnologii. 2020. №4. Str. 85-98. URL: https://cyberleninka.ru/article/n/issledovanie-primeneniya-tehnologii-deception-dlya-predotvrascheniya-ugroz-kiberbezopasnosti (data obrashcheniya: 03.04.2024).

2. Afanas'yeva M.V. Primeneniye dannykh HONEYPOT-sistem dlya preskriptivnoy analitiki deystviy zloumyshlennika. // Aktual'nyye problemy sovremennoy nauki, tekhniki i obrazovaniya. 2023.№1. S.417

3. Shmatova Ye. S. Vybor strategii lozhnoy informatsionnyy sistemy na osnove modeli teorii igr // Voprosy kiberbezopasnosti. 2015. №5 (13). URL: https://cyberleninka.ru/article/n/vybor-strategii-lozhnoy-informatsionnyy-sistemy-na-osnove-modeli-teorii-igr (data obrashcheniya: 13.04.2024).

4. Konyukhovskiy P.V., Shabalin A.A. Teoretiko-igrovyye podkhody v analize strategiy zashchity korporativnykh informatsionnykh sistem // International Journal of Open Information Technologies. 2023 №12. URL: https://cyberleninka.ru/article/n/teoretiko-igrovye-podhody-v-analize-strategiy-zaschity-korporativnyh-informatsionnyh-sistem (data obrashcheniya: 12.04.2024).

5. Konyukhovskiy P.V., Shabalin A.A Game-theoretic Approaches in The Analysis of Corporate Information Systems Protection Strategies // International Journal of Open Information Technologies. 2023 №12. URL: https://cyberleninka.ru/article/n/teoretiko-igrovye-podhody-v-analize-strategiy-zaschity-korporativnyh-informatsionnyh-sistem (date of application: 12.04.2024).

6. J. Pawlick and Q. Zhu, "Deception by design: evidence-based signaling games for network defense," 2015. URL: https://archive.org/details/arxiv-1503.05458/page/n23/mode/2up (date of application: 02.04.2024).

7. Golduyeva D.A., Mokshanina M. A. Prinyatiye resheniy v usloviyakh neopredelennosti na osnove teorii igr s prirodoy. // Vestnik PenzGU. 2020. №4 (32). URL: https://cyberleninka.ru/article/n/prinyatie-resheniy-v-uslovii-neopredelennosti-na-osnove-teorii-igr-s-prirodoy (data obrashcheniya: 14.04.2024).

КОТЕЛЬНИКОВ Никита Дмитриевич, студент 5 курса по специальности «Информационная безопасность автоматизированных систем», ФГБОУ ВО «Магнитогорский государственный технический университет им. Г.И. Носова». 455000, г. Магнитогорск, пр. Ленина 38. E-mail: nick-kotelnik@yandex.ru

АФАНАСЬЕВА Маргарита Владимировна, ассистент кафедры «Информатики и информационной безопасности», ФГБОУ ВО «Магнитогорский государственный технический университет им. Г.И. Носова». 455000, г. Магнитогорск, пр. Ленина 38. E-mail: nansy_stokli@mail.ru

БАРАНКОВА Инна Ильинична, доктор технических наук, заведующий кафедрой «Информатики и информационной безопасности», ФГБОУ ВО «Магнитогорский государственный технический университет им. Г.И. Носова». 455000, г. Магнитогорск, пр. Ленина, 38. E-mail: inna_barankova@mail.ru

KOTELNIKOV Nikita Dmitrievitch, 5th year student majoring in "Information Security of Automated Systems", Federal State Budgetary Educational Institution of Higher Education, Magnitogorsk State Technical University named after G.I. Nosov. 455000, Magnitogorsk, Lenin Ave. 38. E-mail: nick-kotelnik@yandex.ru

AFANASYEVA Margarita Vladimirovna, Assistant of the Department of Computer Science and Information Security, Federal State Budgetary Educational Institution of Higher Education, Magnitogorsk State Technical University named after G.I. Nosov. 455000, Magnitogorsk, Lenin Ave. 38. E-mail: nansy_stokli@mail.ru

BARANKOVA Inna Ilyinichna, Doctor of Technical Sciences, Head of the Department of Computer Science and Information Security, Federal State Budgetary Educational Institution of Higher Education, Magnitogorsk State Technical University named after G.I. Nosov. 455000, Magnitogorsk, Lenin Ave. 38. E-mail: inna_barankova@mail.ru

УДК 004.77 Вестник УрФО № 2(52) / 2024, с. 46–55

Кузьмина У. В., Бачурин И. В., Михайлова О. Е. DOI: 10.14529/secur240205

ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ ЗАЩИТЫ ТЕХНОЛОГИЧЕСКИХ СЕТЕЙ ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

В статье рассмотрены основы построения технологических сетей промышленных предприятий и подходы к обеспечению их защиты. Рассмотрена обобщенная структурная схема АСУ ТП промышленного предприятия. В статье представлен анализ статистики кибератак за 2023 год. Полигоном исследования послужили промышленные предприятия региона. На основе экспериментальных исследований по широкому спектру показателей безопасности получена статистика мер по обеспечению безопасности промышленных предприятий региона. На основе полученной статистики были разработаны рекомендации по обеспечению мер информационной безопасности промышленных сетей предприятий.

Ключевые слова: *информационная безопасность, кибератака, технологические сети, АСУ ТП.Kotelnikov N. D., Afanasyeva M. V., Barankova I. I.*

Kuzmina U. V., Bachurin I. V., Mikhaylova O. E.

PROBLEMS OF PROTECTING TECHNOLOGICAL INDUSTRIAL ENTERPRISES NETWORKS

The article discusses the basics of constructing technological net-works of industrial enterprises and approaches to ensuring their protection. A generalized block diagram of the automated process control system of an industrial enterprise is considered. The article provides an analysis of cyber attack statistics for 2023. The testing ground for the study was the industrial enterprises of the region. Based on experimental studies on a wide range of security indicators, statistics on measures to ensure the security of industrial enterprises in the region were obtained. Based on the statistics obtained, recommendations were developed to ensure information security measures for industrial networks of enterprises.

Keywords: *information security, cyber-attack, technological networks, auto-mated process control system.*

Введение.

Особенности построения архитектуры технологических сетей промышленных предприятий кардинально отличают ее от корпоративных ин-формационных систем: начиная от специфических протоколов передачи данных (Modbus, DP, FDL, FMS), используемого оборудования (датчики, программируемые логические контроллеры, OPC сервера и др.) и программного обеспечения (SCADA, MES системы), заканчивая средой, в которой они функционируют (цеха, производственные помещения). Плюсом к вышеперечисленному в рамках Индустрии 4.0 активно распространяется промышленный интернет вещей (IIoT) и все большее количество устройств подключается к инфраструктуре промышленной сети предприятий.

АСУ ТП строится как децентрализованная система, выполняющая ин-формационные, управляющие и вспомогательные функции и состоит из трех уровней [1]:

Уровень 1 – «полевой» уровень; датчики, исполнительные механизмы, локальные микропроцессорные системы.

Уровень 2 – автоматическое управление; программируемые логические контроллеры (далее ПЛК), получающие данные с «полевого» уровня, передающие данные на верхний уровень для принятия решения по управлению объектом и (или) процессом и формирующие управляющие команды для исполнительных устройств, а также промышленная сеть передачи данных;

Уровень 3 – операторское управление; управление производством; операторские (диспетчерские), инженерные автоматизированные рабочие места, промышленные серверы (SCADA-серверы) с установленным общесистемным и прикладным программным обеспечением, телекоммуникационное оборудование (коммутаторы, маршрутизаторы, межсетевые экраны, иное оборудование), а также каналы связи; формирование, выдача производственно-технологической информации, обмен информацией между цехом и предприятием, прием информации о производственном задании.

Ввод и обмен данными между подсистемами АСУ ТП осуществляется автоматически в режиме реального времени, при этом оператор-технолог осуществляет контроль функционирования системы и участвует в управлении. Режим работы круглосуточный, непрерывный (за исключением запланированных ремонтных работ). В корпоративных информационных системах основной защищаемый ресурс – информация, а цель – обеспечение конфиденциальности. В технологических системах первостепенной задачей является сохранение непрерывности производства, которую обеспечивают доступность и целостности данных. АСУ ТП имеют жестко фиксированную конфигурацию, не допускающую существенных изменений (обновления ПО, использование наложенных средств защиты, корректировка настроек «по умолчанию»).

Актуальность проблемы.

На ряду с ростом устройств, подключаемых к сети, растет и количество хакерских атак на сети. По словам директора по развитию бизнеса в центре противодействия цифровым угрозам Solar JSOC Алексея Павлова: «На сегодня кибербезопасность - один из ключевых вопросов российских компаний». Атакам массово подвергаются крупные предприятия из топливно-энергетического сектора, финансовой отрасли, телеком-индустрии. Среди них - «Газпром», «Лукойл», «Норникель», «Сибур», Сбербанк и др.

В 2023 году в России и СНГ с наибольшим числом кибератак столкнулись организации в сферах промышленности (20% от общего количества инцидентов в регионе), финансов (17%) и ИТ (8%). Таковы данные ежегодного аналитического отчёта, основанного на статистике инцидентов, выявленных у пользователей Kaspersky Managed Detection and Response — решения по круглосуточному мониторингу, проактивному поиску и устранению киберугроз. Российские компании осознают необходимость защиты от целевых атак, поскольку их становится все больше, а последствия обходятся бизнесу очень дорого. Эта тенденция прослеживается не только на российском рынке: в опросе Cybersecurity Insiders 85% организаций-респондентов также не исключают возможности такой атаки на свою инфраструктуру в ближайшие 12 месяцев.

Статистика компании Positive Technologies так же подтверждает постоян-ный рост кибератак (рис. 1). Из них атаки на промышленность составляют 10% (рис 2.).

В связи с тем, что архитектура технологических сетей промышленных предприятий имеет очень сложную многоуровневую структуру, то и обеспече-ние ее защиты является сложной многоуровневой задачей. Для

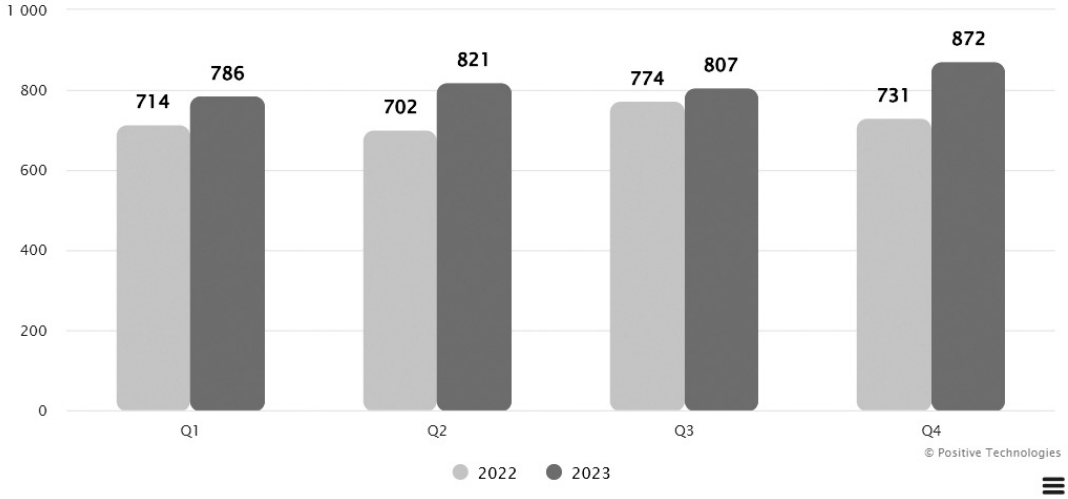

Рис. 1. Количество атак в 2022 и 2023 годах (по кварталам)

решения этой задачи необходимо учитывать требования законодательства для обеспечения защиты информации на критических информационных инфраструктурах [2-7].

Постановка задачи.

Объектом нашего исследования является автоматизированные системы управления технологическими процессами (АСУ ТП) промышленных предприятий региона [8-13].

Обобщенно сетевая инфраструктура АСУ ТП промышленных предпри-ятий имеет 3-уровневую иерархическую топологию сети передачи данных: яд-ро, уровень распределения, уровень доступа. Ядро сети отвечает за высокоскоростную передачу сетевого трафика. Каждое устройство уровня ядра обладает возможностью доступа к любому устройству пункта назначения сети. Устройства уровня ядра соединены между собой отдельными. На уровне распределения происходит суммирование маршрутов и агрегация трафика. В существующем решении уровни распределения и ядра совмещены в одном устройстве. Уровень доступа отвечает за формирование сетевого трафика, выполняет контроль точек входа в сеть и предоставляет службы пограничных устройств. Компоненты АСУ ТП находятся в общей сети предприятия и подключены к автоматизированной системе управления производством (АСУП) с применением меж-

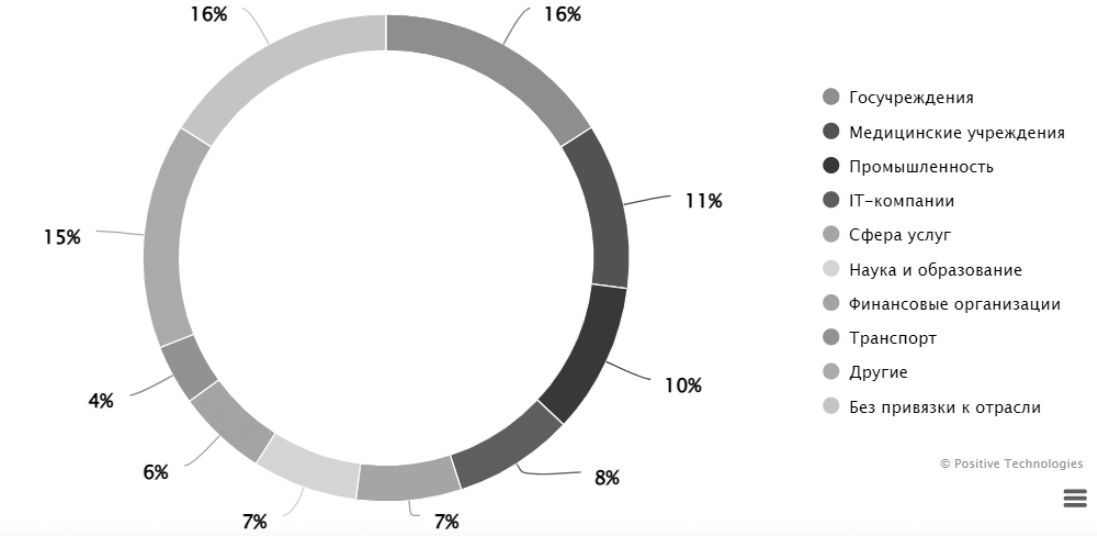

Рис. 2. Категории жертв среди организаций

сетевого экрана. Сеть АСУ ТП не имеет непосредственного соединения с сетью АСУП, для соответствующего взаимодействия два сервера имеют отдельные сетевые интерфейсы, подключенные в соответствующие сети.

Методы оценки защиты технологических сетей промышленных предприятий.

При обеспечении кибербезопасности технологической сети промышленного предприятия одинаково важны, как процессы разработки и реализации защитных мер, так и процессы проверок и контроля состояния защищенности [10]. Подобный контроль дает возможность провести проверку для установления валидности и актуальности используемых средств и систем защиты информации (СЗИ) [9]. На практике у большинства промышленных предприятий нет цельной, четко отлаженной СЗИ. Так, например, антивирусные программные средства установлены на многих АСУ ТП, чего не скажешь о системах обнаружения/предотвращения вторжений или правилах и регламентах реагирования на компьютерные инциденты. Вследствие чего возникает необходимость оценить положение дел и разобраться, какие меры по защите информации реализованы, а какие в обязательном порядке требуют немедленного внедрения. Аудит информационной безопасности (далее аудит ИБ) способствует получению наиболее точных данных о текущем состоянии предприятия в сфере обеспечения безопасности информации [9]. Своевременное обнаружение всех возможных актуальных уязвимостей и угроз безопасности, которые могут возникнуть из-за недостатка принятых мер защищенности, позволит обеспечить построение адекватной и эффективной СЗИ, которая будет соответствовать специфике предприятия.

Аудит ИБ занимает особое положение среди процессов контроля и про-верки, т.к. на данный момент для него не существует строгого нормативного определения. Согласно ГОСТ Р ИСО 19011–2021 [14] «аудит (audit): Система-тический, независимый и документированный процесс установления объектив-ного свидетельства и его объективного оценивания для получения степени соответствия критериям аудита» [4]. В области ИБ принято выделять четыре вида аудита такие как:

1. Экспертный направлен на выявление недостатков СЗИ с помощью экспертов по обеспечению безопасности информации (ОБИ);

2. Оценка соответствия требованиям российского и международного законодательства. Цель настоящего аудита – выявление недостатков СЗИ посредством анализа полноты исполнения требований по ОБИ регламентов, нормативно правовых актов и законодательства;

3. Инструментальный анализ. Данный вид предполагает выявление уяз-вимостей программного и программно-аппаратного обеспечения исследуемой системы;

4. Комплексный аудит включает в себя все вышеперечисленные виды проведения проверки [5].

Международный стандарт ISO 19011-2021 содержит общее представление о процессе аудита ИБ – термины, принципы, этапы и способ оценки компетентности аудитора. Руководствуясь данным документом, аудитор может грамотно и полно разработать программу аудита ИБ и все необходимые организационно-распорядительные документы (далее ОРД), список и содержание которых, от первого этапа «инициирования аудита» и до седьмого «завершение аудита», так же описаны в стандарте. Данные рекомендации применимы для аудита ИБ любых информационных систем, в том числе объектов КИИ. Конкретно для объектов КИИ ФСТЭК России разработал Приказы № 31 [5] и № 239 [2]. Данные документы необходимы для проведения выбранного вида аудита, т. к. содержат базовые наборы требований по «обеспечению защиты информации в автоматизированных системах управления производством и технологическим процессом» (АСУП и АСУ ТП). Кратко можно выделить основные вопросы проведения аудита, которые безусловно потребуют ответов [9]:

1. Какие силы обеспечения ИБ организованы на предприятии?

2. Какие ОРД по ОБИ (и в каком объеме/составе) разработаны и внедрены на предприятии?

3. Какие внедрены программные/ программно-аппаратные СЗИ и каков срок действия их сертификатов?

4. Какие осуществляются, как реализованы и чем регламентированы мероприятия для обеспечения безопасности информации?

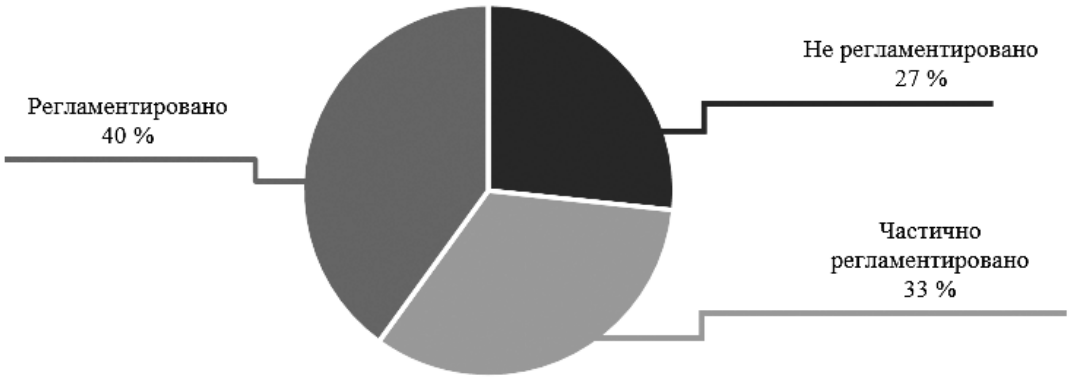

Рис. 3. Анализ полноты ОРД по ОБИ промышленных предприятий региона

Результаты исследования защищенности технологических сетей промышленных предприятий.

Выполнение требований 239-го приказа [2] необходимо для объектов критической информационной инфраструктуры (ОКИИ), признанных значимыми на основании проведенной процедуры категорирования, по правилам, утвержденным Постановлением Правительства РФ № 127 [6]. Незначимые объекты КИИ должны выполнять требования 31-го приказа [5], а также обязанности ч.2 ст. 9 Федерального закона № 187 [4] (требования данного ФЗ распространяются и на значимые объекты). Наименование мер защиты информации в обоих приказах идентичны, их различие состоит в том, что для незначимых объектов перечень мер определен для каждого из уровней значимости обрабатываемой в них информации, в то время как для значимых – по трем категориям значимости.

На основе проведенных исследований предприятий региона получена следующая статистика обеспечения безопасности промышленных предприятий региона (рис. 3, рис. 4).

В нормативных документах разработан базовый набор мер, которые в обязательном порядке должны выполняться для обеспечения защиты информации на предприятии. На диаграмме (рис. 4) представлен анализ полноты реализации базового состава технических мер КИИ, согласно нормативным документам. Обозначения мер на диаграмме (рис. 4) приведены согласно 239-му приказу [2].

Нулевой уровень показывает, что мера безопасности ничем не регламен-тирована и не реализована технически, первый уровень – разработаны и вне-дрены организационные меры защиты информации, второй – внедрены техни-ческие средства защиты информации, третий – разработаны и внедрены как организационные, так и технические меры защиты информации. Из диаграммы видно, например, что ИПО.0 (Регламентация правил и процедур информирования и обучения персонала), ИПО.1 (Информирование персонала об угрозах безопасности информации и о правилах безопасной работы), ИПО.2 (Обучение персонала правилам безопасной работы, ИПО.3 (Проведение практических занятий с персоналом по правилам безопасной работы), ИПО.4 (Контроль осведомленности персонала об угрозах безопасности информации и о правилах безопасной работы) имеют нулевой уровень. Следовательно, по этим показателям либо полностью отсутствует, либо выполнены точечно требования безопасности. То же самое касается и показателей

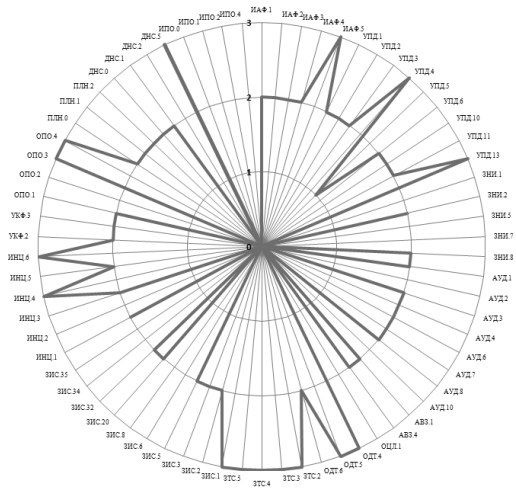

Рис. 4. Анализ полноты применения технических мер по ОБИ

ЗИС.35 (Управление сетевыми соединениями), ЗИС.36 (Создание (эмуляция) ложных компонентов автоматизированных систем), ЗИС.6 (Управление сетевыми потоками), ЗИС.8 (Сокрытие архитектуры и конфигурации автоматизированной системы) и т. д.

Заключение.

Проведенное исследование ИБ показало, что текущее со-стояние АСУ ТП в сфере обеспечения безопасности информации, не в полной мере удовлетворяет требованиям Приказов № 235, 239 ФСТЭК России. По результатам исследования ИБ составлены перечни организационных и технических мер, применение которых позволит перекрыть необходимый базовый набор мер обеспечения безопасности, регламентированный Приказами ФСТЭК России. Базовый набор мер защиты информации необходимо адаптировать в соответствии с применяемыми информационными технологиями и особенностями функционирования значимого объекта. Меры: контроль доступа из внешних систем, антивирусная защита электронной почты и иных сервисов, защита беспроводных соединений, управление перемещением виртуальных машин и обрабатываемых на них данных - следует исключить, т. к. они не применимы ввиду отсутствия технологий. В адаптивный набор необходимо включить меры группы «Предотвращение вторжений»: регламентация правил и процедур предотвращения вторжений, обнаружение и предотвращение компьютерных атак, обновление базы решающих правил. В результате исследования предложены общие рекомендации по реализации организационных и технических мер защиты информации на промышленных предприятиях.

Управление рисками ИБ АСУ ТП должно базироваться на требованиях ГОСТ Р ИСО/МЭК 27005. Управление рисками ИБ АСУ ТП должно включать:

а) инвентаризацию и классификацию активов АСУ ТП, включая инфор-мацию, обрабатываемую и хранящуюся в АСУ ТП;

б) оценку эффективности существующих мер и контроля по ИБ, включая выявление недостатков и уязвимостей обеспечения ИБ АСУ ТП, а также оценку рисков ИБ АСУ ТП;

в) планирование процессов и мер по ИБ АСУ ТП, направленных на снижение рисков, в том числе внесение изменений в существующие процессы и меры;

г) реализацию планируемых процессов и мер, осуществление контроля за выполнением планов и оценку достигнутых результатов.

Оценка рисков в отношении каждой из АСУ ТП должна осуществляться:

а) на плановой основе для АСУ ТП - не реже одного раза в три года;

б) постоянно с учетом изменений в технологических процессах, архитектуре и инфраструктуре (вычислительной, сетевой, КИПиА) АСУ ТП, а также по результатам мониторинга, аудитов ИБ АСУ ТП, инцидентов ИБ и оценки эффективности процессов и мер по обеспечению ИБ АСУ ТП.

Ответственность за проведение и координацию работ, связанных с управлением рисками нарушения ИБ АСУ ТП, включая утверждение результатов оценки рисков, утверждение планируемых мер по снижению рисков и выделение требуемых ресурсов на реализацию этих мер, возлагается на руководителя, отвечающего за обеспечение ИБ АСУ ТП.

Актуальность сведений об АСУ ТП должна поддерживаться путем проведения инвентаризации не реже одного раза в год. Ответственность за организацию регулярной инвентаризации должен нести руководитель, ответственный за обеспечение ИБ АСУ ТП.

В случае изменений в составе оборудования или ПО, сетевой или вычис-лительной инфраструктуры АСУ ТП эксплуатационный персонал АСУ ТП должен передать сведения об этих изменениях для актуализации информации в паспорте ИБ АСУ ТП.

Согласно Стандарту, сеть АСУ ТП должна быть разделена на «зоны». Нарушение связи между зонами АСУ ТП не должно приводить к остановке технологического процесса, он должен продолжаться, в режиме ограниченной функциональности, до восстановления связей между зонами.

Выделяют пять зон:

1) Зона систем управления содержит системы управления технологи-ческими процессами и представляет собой систему, с архитектурой человеко-машинного интерфейса – SCADA.

2) Зона систем промышленной безопасности включает системы противоаварийной защиты (далее - ПАЗ), обеспечивающих распознавание отклонения процесса от заданных параметров и оповещение об аварийных ситуациях, а также автоматическую остановку технологического процесса в случае обна-

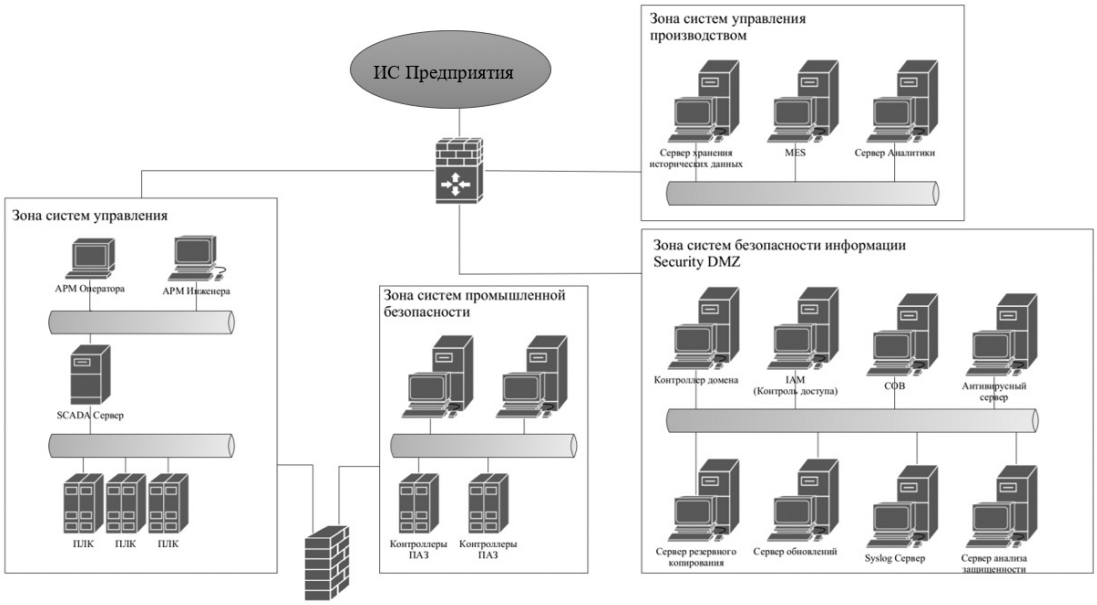

Рис. 5. Пример практической реализации концепции МЭК 62443 на обобщенной структурной схеме АСУ ТП промышленного предприятия

руженной критической неисправности. Данная зона должна быть максимально изолирована и защищена в соответствии с концепцией «защита в глубину».

3) Зона систем управления производством. К данной зоне, как прави-ло, относят системы, размещенные на уровне выше, чем SCADA, и не оказы-вающие прямое влияние на ход технологического процесса. Это могут быть системы класса Historian, системы оперативного управления производством. Например, MES, системы класса Intelligence и другие.

4) Зона управления СЗИ располагаются технические, программные и программно-аппаратные средства, обеспечивающие защиту 1-3 зон от угроз безопасности информации. Такими средствами могут быть: сервера - сбора и анализа данных, антивирусной защиты, резервного копирования; контроллер домена, системы обнаружения вторжений, и т.д.

5) Демилитаризованная зона (далее - ДМЗ) содержит технические или программные средства, которые обеспечивают взаимодействия между выше-описанными зонами АСУ ТП и корпоративной сетью предприятия (рис. 5).

Подсистема защиты периметра должна обеспечивать защиту информации от несанкционированного доступа, сегментирование технологической сети и контролировать межсетевое взаимодействие между сегментами технологической сети, технологическими ДМЗ и корпоративная информационная структура (КИС).

Литература

1. Информационное сообщение ФСТЭК России, Об утверждении ме-тодического документа ФСТЭК России «Методика оценки показателя состоя-ния защиты информации и обеспечения безопасности объектов критической информационной инфраструктуры Российской Федерации». [Электронный ре-сурс] Режим доступа: https://fstec.ru/dokumenty/vse-dokumenty/informatsionnye-i-analiticheskie-materialy/informatsionnoe-soobshchenie-fstek-rossii

2. Приказ ФСТЭК России от 25 декабря 2017 г. № 239 «Об утверждении требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации» [Электронный ресурс] Режим доступа: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239

3. Приказ ФСТЭК России от 18 февраля 2013 г. № 235 «Об утверждении требований к созданию систем безопасности значимых объектов критической информационной инфраструктуры Российской Федерации и обеспечению их функционирования» [Электронный ресурс] Режим доступа: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-21-dekabrya-2017-g-n-235

4. Федеральный закон от 19 июля 2017 г. № 187 «О безопасности критической информационной инфраструктуры Российской Федерации» [Электронный ресурс] Режим доступа: https://fstec.ru/dokumenty/vse-dokumenty/zakony/federalnyj-zakon-ot-26-iyulya-2017-g-n-187-fz

5. Приказ ФСТЭК России от 14 марта 2014 г. № 31 «Об утверждении требований к обеспечению защиты информации в автоматизированных систе-мах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды» [Электронный ресурс] Режим доступа: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-14-marta-2014-g-n-31

6. Постановление Правительства от 8 февраля 2018 г. № 127 «Об утверждении правил категорирования объектов критической информационной инфраструктуры российской федерации, а также перечня показателей критериев значимости объектов критической информационной инфраструктуры российской федерации и их значений» » [Электронный ресурс] Режим доступа: https: //fstec.ru/tekhnicheskaya-zashchita-informatsii/obespechenie-bezopasnosti-kii/287-postanovleniya/1614-postanovlenie-pravitelstva-rossijskoj-federatsii-ot-8-fevralya-2018-g-n-127

7. КИИ: обзор нормативной базы ФСТЭК России [Электронный ресурс] Режим доступа: https://ics-cert.kaspersky.ru/publications/reports/2018/06/25/obzor-normativnoy-bazy-fstek-rossii/

8. Афанасьева С.В., Кузьмина У.В. / Основные проблемы при работе с центрами мониторинга информационной безопасности //\Вестник УрФО. Безопасность в информационной сфере. 2023. № 1 (47). С. 51-58.

9. Баранкова И.И., Семавина Е.А., Михайлова У.В. / Аудит информа-ционной безопасности промышленных предприятий, направленный на оценку соответствия требованиям российского и международного законодательства / Вестник УрФО. Безопасность в информационной сфере. 2022. № 3 (45). С. 76-82.

10. Михайлова У.В., Быкова Т.В. / Аудит информационной безопасности предприятия ООО "Машиностроительный завод РИВС" // Актуальные проблемы современной науки, техники и образования. Тезисы докладов 77-й международной научно-технической конференции. 2019. С. 416-417.

11. Михайлова У.В., Афанасьева М.В. / Аудит информационной безо-пасности предприятия ООО "АНСЕР" // Актуальные проблемы современной науки, техники и образования. Тезисы докладов 77-й международной научно-технической конференции. 2019. С. 417-418.

12. Баранкова И.И., Михайлова У.В., Афанасьева М.В., Афанасьев М.Ю. / Принципы построения модели надежности системы защиты информации АСУ ТП доменной печи // Актуальные проблемы современной науки, техники и образования. Тезисы докладов 77-й международной научно-технической конференции. 2019. С. 424.

13. Barankova, I.I., Mikhailova, U.V., Kalugina, O.B. / Analysis of the Problems of Industrial Enterprises Information Security Audit // Lecture Notes in Electrical EngineeringЭта ссылка отключена., 2020, 641 LNEE, p. 976–985.

14. Национальный стандарт Российской Федерации. ГОСТ Р ИСО 19011-2021 «Оценка соответствия. Руководящие указания по проведению аудита систем менеджмента» [Текст], Принят Приказом Федерального агентства по техническому регулированию и метрологии от 21 апреля 2021 г. – 2021. – 41 с

References

15. Informatsionnoye soobshcheniye FSTEK Rossii, Ob utverzhdenii me-todicheskogo dokumenta FSTEK Rossii «Metodika otsenki pokazatelya sostoya-niya zashchity informatsii i obespecheniya bezopasnosti ob"yektov kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii». [Elektronnyy re-surs] Rezhim dostupa: https://fstec.ru/dokumenty/vse-dokumenty/informatsionnye-i-analiticheskie-materialy/informatsionnoe-soobshchenie-fstek-rossii

16. Prikaz FSTEK Rossii ot 25 dekabrya 2017 g. № 239 «Ob utverzhde-nii trebovaniy po obespecheniyu bezopasnosti znachimykh ob"yektov kritiche-skoy informatsionnoy infrastruktury Rossiyskoy Federatsii» [Elektron-nyy resurs] Rezhim dostupa: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239

17. Prikaz FSTEK Rossii ot 18 fevralya 2013 g. № 235 «Ob utverzhde-nii trebovaniy k sozdaniyu sistem bezopasnosti znachimykh ob"yektov kritiche-skoy informatsionnoy infrastruktury Rossiyskoy Federatsii i obespecheniyu ikh funktsionirovaniya» [Elektronnyy resurs] Rezhim dostupa: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-21-dekabrya-2017-g-n-235

18. Federal'nyy zakon ot 19 iyulya 2017 g. № 187 «O bezopasnosti kri-ticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii» [Elek-tronnyy resurs] Rezhim dostupa: https://fstec.ru/dokumenty/vse-dokumenty/zakony/federalnyj-zakon-ot-26-iyulya-2017-g-n-187-fz

19. Prikaz FSTEK Rossii ot 14 marta 2014 g. № 31 «Ob utverzhdenii trebovaniy k obespecheniyu zashchity informatsii v avtomatizirovannykh siste-makh upravleniya proizvodstvennymi i tekhnologicheskimi protsessami na kri-ticheski vazhnykh ob"yektakh, potentsial'no opasnykh ob"yektakh, a takzhe ob"yektakh, predstavlyayushchikh povyshennuyu opasnost' dlya zhizni i zdorov'ya lyudey i dlya okruzhayushchey prirodnoy sredy» [Elektronnyy resurs] Rezhim dostupa: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-14-marta-2014-g-n-31

20. Postanovleniye Pravitel'stva ot 8 fevralya 2018 g. № 127 «Ob utverzhdenii pravil kategorirovaniya ob"yektov kriticheskoy informatsionnoy infrastruktury rossiyskoy federatsii, a takzhe perechnya pokazateley kriteri-yev znachimosti ob"yektov kriticheskoy informatsionnoy infrastruktury ros-siyskoy federatsii i ikh znacheniy» » [Elektronnyy resurs] Rezhim dostupa: https: //fstec.ru/tekhnicheskaya-zashchita-informatsii/obespechenie-bezopasnosti-kii/287-postanovleniya/1614-postanovlenie-pravitelstva-rossijskoj-federatsii-ot-8-fevralya-2018-g-n-127

21. KII: obzor normativnoy bazy FSTEK Rossii [Elektronnyy re-surs] Rezhim dostupa: https://ics-cert.kaspersky.ru/publications/reports/2018/06/25/obzor-normativnoy-bazy-fstek-rossii/

22. Afanas'yeva S.V., Kuz'mina U.V. / Osnovnyye problemy pri rabote s tsentrami monitoringa informatsionnoy bezopasnosti //\Vestnik UrFO. Bez-opasnost' v informatsionnoy sfere. 2023. № 1 (47). S. 51-58.

23. Barankova I.I., Semavina Ye.A., Mikhaylova U.V. / Audit informa-tsionnoy bezopasnosti promyshlennykh predpriyatiy, napravlennyy na otsenku sootvetstviya trebovaniyam rossiyskogo i mezhdunarodnogo zakonodatel'stva / Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. 2022. № 3 (45). S. 76-82.

24. Mikhaylova U.V., Bykova T.V. / Audit informatsionnoy bezopasno-sti predpriyatiya OOO "Mashinostroitel'nyy zavod RIVS" // Aktual'nyye problemy sovremennoy nauki, tekhniki i obrazovaniya. Tezisy dokladov 77-y mezhdunarodnoy nauchno-tekhnicheskoy konferentsii. 2019. S. 416-417.

25. Mikhaylova U.V., Afanas'yeva M.V. / Audit informatsionnoy bez-opasnosti predpriyatiya OOO "ANSER" // Aktual'nyye problemy sovremennoy nauki, tekhniki i obrazovaniya. Tezisy dokladov 77-y mezhdunarodnoy nauchno-tekhnicheskoy konferentsii. 2019. S. 417-418.

26. Barankova I.I., Mikhaylova U.V., Afanas'yeva M.V., Afanas'yev M.YU. / Printsipy postroyeniya modeli nadezhnosti sistemy zashchity informa-tsii ASU TP domennoy pechi // Aktual'nyye problemy sovremennoy nauki, tekh-niki i obrazovaniya. Tezisy dokladov 77-y mezhdunarodnoy nauchno-tekhnicheskoy konferentsii. 2019. S. 424.

27. Barankova, I.I., Mikhailova, U.V., Kalugina, O.B. / Analysis of the Prob-lems of Industrial Enterprises Information Security Audit // Lecture Notes in Electri-cal EngineeringЭта ссылка отключена., 2020, 641 LNEE, p. 976–985.

28. Natsional'nyy standart Rossiyskoy Federatsii. GOST R ISO 19011-2021 «Otsenka sootvetstviya. Rukovodyashchiye ukazaniya po provedeniyu audita sistem menedzhmenta» [Tekst], Prinyat Prikazom Federal'nogo agentstva po tekhnicheskomu regulirovaniyu i metrologii ot 21 aprelya 2021 g. – 2021. – 41 s

КУЗЬМИНА Ульяна Владимировна, кандидат технических наук, до-цент кафедры «Информатики и информационной безопасности», ФГБОУ ВО «Магнитогорский государственный технический университет им. Г.И. Носова». 455000, г. Магнитогорск, пр. Ленина, 38. E-mail: u.mihaylova@magtu.ru

БАЧУРИН Иван Владимирович, аспирант 2 курса кафедры «Вычислительных машин, систем и сетей», ФГБОУ ВО «Национальный исследовательский университет «МЭИ». 111250, г. Москва, Красноказарменная улица, дом 14, стр. 1. E-mail: biv@ipc2u.ru

МИХАЙЛОВА Ольга Евгеньевна, студент кафедры «Информатики и информационной безопасности», ФГБОУ ВО «Магнитогорский государствен-ный технический университет им. Г.И. Носова». 455000, г. Магнитогорск, пр. Ленина, 38. E-mail: olgamihailova01@mail.com

KUZMINA Ulyana Vladimirovna, Candidate of Technical Sciences, Associate Professor of the student of the Department of Computer Science and Information Security, Federal State Budgetary Educational Institution of Higher Education, Magnitogorsk State Technical University named after G.I. Nosov. 455000, Magnitogorsk, Lenin Ave. 38. E-mail: u.mihaylova@magtu.ru

BACHURIN Ivan Vladimirovich, 2nd year graduate student of the Depart-ment of Computer Machines, Systems and Networks, Federal State Budgetary Educational Institution of Higher Education "National Research University "MPEI". 111250, Moscow, Krasnokazarmennaya street, house 14, building 1. E-mail: biv@ipc2u.ru

MIKHAYLOVA Olga Evgenevna, student of the Department of Computer Science and Information Security, Federal State Budgetary Educational Institution of Higher Education, Magnitogorsk State Technical University named after G.I. Nosov. 455000, Magnitogorsk, Lenin Ave. 38. E-mail: olgamihailova01@mail.com

УДК 004.056.5 Вестник УрФО № 2(52) / 2024, с. 56–67

Ожгибесова А. С., Шабуров А. С., Южаков А. А. DOI: 10.14529/secur240206

ОБ ОЦЕНКЕ РИСКОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ПРИМЕНЕНИЯ НЕЧЕТКИХ КОГНИТИВНЫХ КАРТ В ИНТЕЛЛЕКТУАЛЬНЫХ ТРАНСПОРТНЫХ СИСТЕМАХ УПРАВЛЕНИЯ ДОРОЖНЫМ ДВИЖЕНИЕМ

В статье проанализирована проблема оценки рисков безопасности информации в условиях внедрения интеллектуальных информационных систем, а также роста количества и сложности кибератак. Приведен состав интеллектуальной транспортной системы, а также математическое описание нечеткой когнитивной карты и последовательность ее разработки. Проанализирована структура информационных потоков одной из транспортных систем и основные задачи ее функционирования. На основе анализа векторов атак на информационную транспортную систему определены необходимые концепты для разработки нечеткой когнитивной карты оценки рисков. Приведена количественная оценка программного обеспечения, с использованием которого разработана модель оценки рисков на основе нечеткой когнитивной модели

Ключевые слова: *риск информационной безопасности, защита информа-ции, интеллектуальная транспортная система, нечеткая когнитивная карта*

ABOUT ASSESSING INFORMATION SECURITY RISKS BASED ON THE APPLICATION OF FUZZY COGNITIVE MAPS IN INTELLIGENT TRANSPORT TRAFFIC MANAGEMENT SYSTEMS

The article analyzes the problem of assessing information security risks in the context of the implementation of intelligent information systems, as well as the growth in the number and complexity of cyberattacks. The composition of an intelligent transport system is presented, as well as a mathematical description of a fuzzy cognitive map and the sequence of its development. The structure of information flows of one of the transport systems and the main tasks of its functioning are analyzed. Based on the analysis of attack vectors on the information transport system, the necessary concepts for developing a fuzzy cognitive risk assessment map are identified. A quantitative as-sessment of the software is provided, using which a risk assessment model has been de-veloped based on a fuzzy cognitive model

Keywords: *information security risk, information protection, intelligent transport system, fuzzy cognitive map*

Введение

На этапе оценки рисков информационной безопасности возникают достаточно сложные проблемы, обусловленные необходимостью принятия оперативных решений[1]. Данные решения требуют адекватного оценивания параметров и характеристик векторов информационных атак, что также обусловленоразнообразием, сложностью и непредсказуемостью как самих атак, так и прочих деструктивных информационных воздействий. Особенно актуальна данная проблема в условиях внедрения и развития интеллектуальных информационных систем, обеспечивающих функционирование различных, в том числе и критически важных процессов и систем[2]. Необходимость оперативного принятия решений по защите информации требует совершенствования процесса оценки рисков, в том числе и за счет автоматизации с использованием компьютерных средств моделирования.

Прикладной характер процесса оценки рисков предполагает формирование необходимой и достаточной инструментальной базы, позволяющей оперативно формировать требуемые модели угроз безопасности информации. Например, внедрение концепции «умного города» предполагает использование подобных автоматизированных решений для инфраструктурных транспортных решений на основе интеллектуальной поддержки[3].

Современные интеллектуальныетранспортныесистемы (ИТС) представляют собой интеграцию информационных и коммуникационных технологий и средств автоматизации с транспортной инфраструктурой, транспортными средствами и их пользователями[4]. Внедрение подобных систем позволяет реализовать с заданным качеством процессы управления транспортной системой региона, населенного пункта, или отдельной дороги, а также группой транспортных средств, или отдельным средством. Основной целью внедрения подобных систем является обеспечения качественного использования дорожной сети, повышения ее безопасности, а также повышение эффективности транспортного процесса. Достижение цели предполагает использование значительного количества про-

граммно-аппаратурных средств, технических, периферийных устройств, каналов связи для осуществления сбора, обработки и передачи значительных объемов данных.

К основным компонентам ИТС, как правило, относятся следующие компоненты и системы[3]:

1) Транспортные средства, оборудованные средствами для коммуникации и управления.

2) Средства связи и коммуникации, включающие системы передачи данных, электронные сайты, приложения и сети.

3) Системы мониторинга дорожной ситуации, включающие оконечные устройства для мониторинга: видеокамеры, датчики, придорожные метеостанции.

4) Системы общего управления, включающие средства контроля дорожного трафика, паркинга, управления ремонтом и строительством дорожной инфраструктуры, средства управления общественным транспортом.

5) Системы управления движением, включающие средства управления светофорными объектами, шлагбаумами, электронными дорожными знаками.

6) Платежные системы и приложения, включающие платежные автоматы, системы продажи электронных билетов, средства бесконтактной идентификации (RFID-метки).

Сложность взаимодействия и реализация функционального предназначения ИТС, неопределенность и непредсказуемость ситуаций транспортного обеспечения обуславливает значительное количество уязвимостей и векторов информационных атак. Последствия подобных атак сложно предсказуемы и в совокупности влекут за собой значительный экономический ущерб, а также возможный ущерб безопасности жизни и здоровья всех участников дорожного движения.

Противодействие подобным информационным атакам на ИТС предполагается на основе внедрения системы защиты информации. В свою очередь система защиты информации, а также эффективность мер по обеспечению информационной безопасности напрямую зависит от оперативности и адекватности оценки рисков, основанной на автоматизированных методах[5]. Одним из проверенных инструментов подобной оценки является использование метода на основе нечетких когнитивных карт [6].

Метод оценки рисков информационной безопасности на основе нечетких когнитивных карт

Как правило, нечеткая когнитивная карта (НКК) – это ориентированный граф, заданный с помощью множеств[6]:

$$\text{НКК}=<C,F,W>,$$

где $C=\{C_i\}$ – множество концептов – вершин графа, которые в данном случае выступают в качестве факторов, наиболее значимых для оценки угроз безопасности информации;

$F=\{F_i\}$ – множество направленных дуг графа – связей между концептами;

$W=\{W_i\}$ – множество весов связей графа, которые могут быть положительными ($W_{ij}>0$) или отрицательными ($W_{ij}<0$) в зависимости от того, усиливают или ослабляют влияние концепта на концепт.

Разработка НКК предполагает реализацию нескольких обязательных этапов, с учетом специфики решаемых задач, а также используемых инструментов моделирования.

На первоначальном этапе определяются цели и задачи моделирования, которые достигаются и решаются посредством разработки НКК.

На втором этапе определяются исходные данные, которые, как правило, носят статистический характер, представляют собой результаты исследований, или экспертных оценок.

На третьем этапе определяются правила вывода результатов моделирования, обусловленные логическими законами принятия решений и основанные на имеющейся базе данных. Правила вывода определяют связи получаемых решений, основанных на множестве входных параметров.

На четвертом этапе разработки НКК предполагается применение специализированного программного обеспечения, с целью проверки работоспособности модели.

На заключительном этапе проводится тестирования и валидация разработанной модели на основе созданных разнообразных сценариев, в меняющихся условиях. Валидация модели осуществляется с учетом экспертных оценок и реальных условий функционирования моделируемых систем. Внедрение разработанной модели предполагает ее применение в предметной области для решения конкретных задач[7].

Выбор программного обеспечения для моделирования НКК

Выбор программного обеспечения для реализации модели напрямую зависит от целей и задач, которые преследуются в работе, а также от финансовых ограничений и технических возможностей.

Проанализировав информацию о существующих программных продуктах, можно выделить следующие наиболее важные критерии выбора[8-11]:

1. Простота использования программы. Интуитивно понятный интерфейс позволит пользователям быстро освоить процесс создания моделей, что в свою очередь приведет к экономии времени оценки угроз информационной безопасности.

2. Возможность визуального представления моделей, что позволяет сделать модели более наглядными добавлению цветовых эффектов, текста и мультимедийных объектов. Таким образом, упрощается проведение анализа и интерпретация полученных результатов.

3. Широкое разнообразие инструментов форматирования и стилизации обеспечивает большую гибкость и творческие возможности при создании моделей, что сделает их более информативными и удобными для восприятия.

4. Возможность экспорта данных в различные форматы для использования полученных исследований в проектах, при добавлении в отчеты и интеграции с другими приложениями.

5. Открытый исходный код, обеспечивающий прозрачность в отношении функциональности и безопасности программного обеспечения. Код можно проверить на наличие ошибок, дополнить или модифицировать в соответствии с потребностями организации.

6. Поддержка разработчиками и периодические обновления (сопровождение) позволяют избавляться от возникающих ошибок функционирования, увеличивать возможности моделирования и совершенствовать продукт.

7. Мультиплатформенность, что делает программу доступной для широкого круга пользователей не зависимо от операционной системы или устройства, которые он использует.

Кроме того, необходимо ввести количественную оценку критериев для получения объективных результатов. Оптимальное решение принимается при точном расчете и сравнении различного программного обеспечения (ПО) на основе одних и тех же мер или показателей[12].

Для удобства представления критериев выбора ПО разработана сводная таблица характеристик программных решений, на основе присвоения им одного из следующих значений:

0 – критерий отсутствует в программе;

1 – критерий присутствует, но функционально ограничен;

2 – критерий присутствует и имеет полную функциональность.

Полученные оценки значений критериев ПО, потенциально используемых для разработки НКК, представлены в таблице 1. Итоговый суммарный результат оценки критериев позволяет осуществить выбор наиболее эффективное ПО для разработки НКК.

Таблица 1.

Количественная оценка программного обеспечения

	FCMapper	CmapTools	XMind	Visual Understanding Environment
Простота использования	1	2	2	2
Возможности визуального представления моделей	1	2	2	2
Разнообразие функций	0	2	2	2
Экспорт данных	1	1	1	2
Открытый исходный код	2	0	0	2
Поддержка и обновления	0	1	2	1
Мультиплатформенность	1	2	2	2
Итоговая оценка	6	10	11	13

Результаты количественной оценки критериев позволяет выбрать ПО Visual Understanding Environment (VUE), являющееся оптимальным для решения поставленных задач текущего исследования.

Разработка нечеткой когнитивной модели для автоматизации оценки рисков

1. Цель и задачи моделирования

Целью моделирования является наглядное представление метода оценки рисков на основе НКК для конкретного примера в организации, управляющей интеллектуальной транспортной системой.

Задачи:

- анализ особенностей организации, актуальных угроз (дестабилизирующих факторов), информационных ресурсов, наиболее значимых рисков, являющихся целевыми факторами;
- построение графа связности концептов и определение их весов;
- расчет относительного уровня риска информационной безопасности (целевых факторов).

2. Определение входных данных

В качестве предметной области в данной работе рассматривается одна из транспортных систем, предназначенная для контроля и обеспечения безопасности дорожного движения, функционирующая на базе государственного краевого учреждения «Центр безопасности дорожного движения» (ГКУ ЦБДД). Оценка рисков осуществляется на примере одного из значимых объектов КИИ. На рис. 1 представлена структура информационных потоков ГКУ «ЦБДД».

По результатам экспертной оценки функционирования систем ГКУ «ЦБДД» определен перечень информационных систем - объектов КИИ, к которым относятся:

1. Информационная система «Метеорологическое обеспечение»;
2. Информационная система «Фотовидеофиксация»;
3. Информационная система «Видеонаблюдение»;
4. Информационная система «Весовой и габаритный контроль».

В качестве примера для построенияНКК определена информационная система (ИС) «Метеорологическое обеспечение», предназначенная для выполнения следующих функций:

- получение оперативной информации о погодных условиях и состоянии дорожного покрытия на сети автомобильных дорог Пермского края;
- прогнозирование возможных опасных метеорологических условий на дорожном полотне;
- принятие решений по проведению необходимых работ по содержанию дорог.

В состав ИС входят комплексные посты дорожного контроля (КПДК), основу которых составляют 59 датчиков, распределенных по дорожной сети региона.

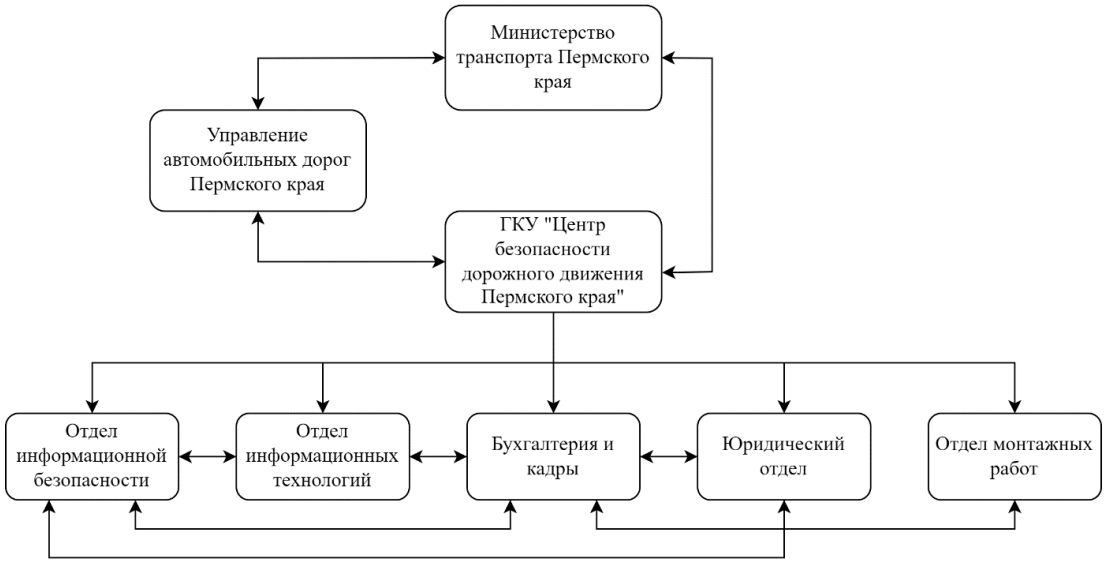

Рисунок 1 – Структура информационных потоков ГКУ «ЦБДД»

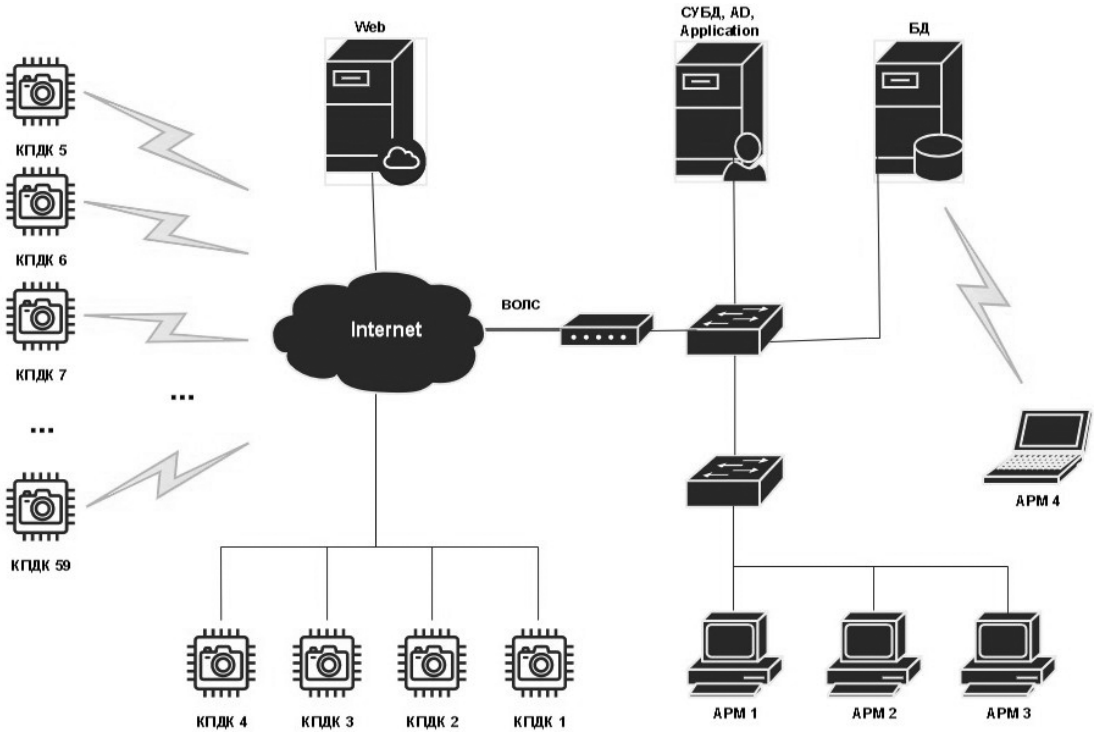

Рис. 2. Общая схема ИС «Метеорологическое обеспечение»

Структурная схема ИС «Метеорологическое обеспечение», включающая 4-е автоматизированных рабочих места (АРМ) операторов ИС, объединенных в одну локальную сеть.

Управление ИС осуществляется на основе 3-х серверных станций, обеспечивающих выполнение всех вышеперечисленных функций.

В ИС «Метеорологическое обеспечение» обрабатываютсяследующие виды информации:

• видовая информация о погодной обстановке на дорожном полотне в зоне обозрения видеокамеры (формируется КПДК во взаимодействии с подключенными к ним камерами видеонаблюдения);

• информация в числовом и текстовом виде о погодных условиях, в которых располагается станция (формируется КПДК во взаимодействии с подключенными к ним датчиками).

Полученная от датчиков информация направляется по каналам связи через КПДК для хранения на сервере баз данных. В дальнейшем данная информация используется оператором диспетчерского центра, отделом эксплуатации автомобильных дорог для принятия решения по управлению дорожным движением.

Основными задачами обработки информации в системе являются: составление статистики погодных условий с целью предварительной оценки затрат на обслуживание дорожного полотна, уведомление подрядных организаций о необходимости принять меры по уходу за дорожным полотном, общедоступное доведение до автовладельцев сведений о погодных условиях на дорожном полотне. Рассматриваемая сеть относится к сетям электросвязи 2 категории (выделенная)[13].

К основным последствиям реализации компьютерных атак (КА) на ИС «Метеорологическое обеспечение» относятся следующие:

• прекращение или нарушение функционирования объектов транспортной инфраструктуры (автомобильных дорог), так как объект обеспечивает осведомление подрядных организаций, ответственных за обеспечение надлежащего состояния дорожного полотна;

• нарушение функционирования Министерства транспорта Пермского края в части выполнения возложенной на него функции по осуществлению государственного контроля (надзора) за соблюдением требований технического регламента Таможенного союза «Безопасность автомобильных дорог».

Анализ деструктивных информационных воздействий на ИС «Метеорологическое обеспечение» предполагающий оценку потенциально возможных векторов КА, которые могут быть реализованы злоумышленником, представлены на рис. 3.

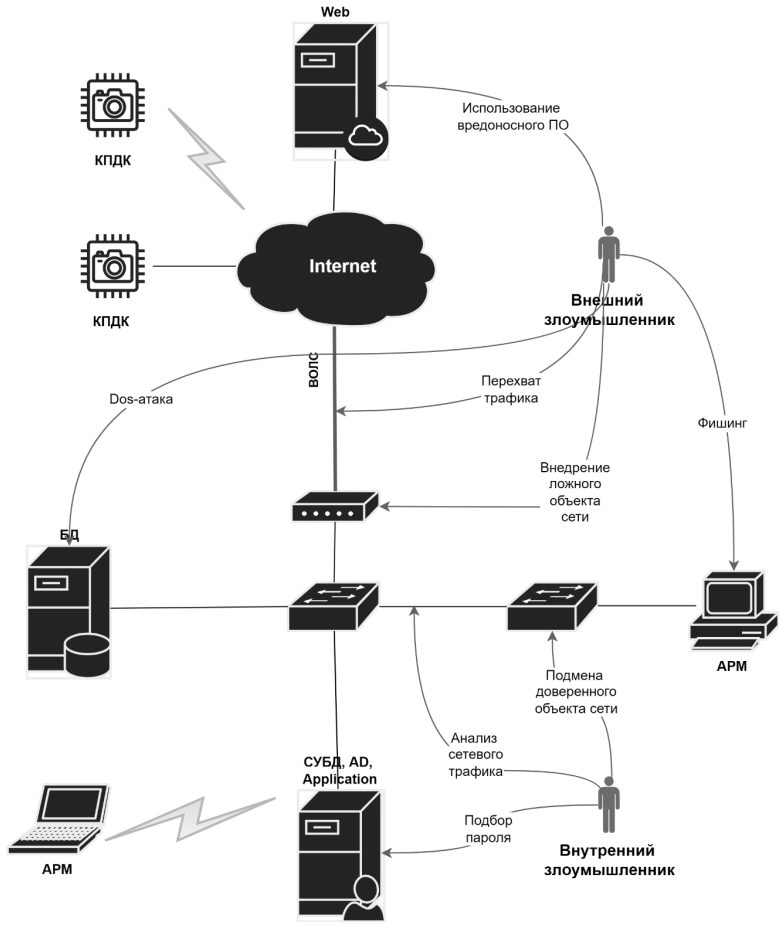

Рис. 3. Вектора атак на ИТС

Внутренний злоумышленник, действующий из корпоративного сегмента сети, может проникнуть в технологическую сеть предприятия и скомпрометировать технологический процесс, получить и повысить привилегий в ОС на узлах ИС.

Внешний злоумышленник, обнаруживший доступные интерфейсы администрирования серверов ИС и удаленного доступа к СУБД в совокупности с повсеместным использованием словарных и стандартных паролей привилегированных пользователей, может в один шаг получить полный контроль как над веб-приложениями, так и над серверами, получить доступ к БД и файлам, развивать атаку на другие ресурсы. Хранящиеся в открытом доступе важные данные, например учетные записи, исходный код веб-приложений, персональные данные пользователей, могут быть использованы при атаках.

Процесс работы ИТС начинается со сбора информации о состоянии трафика. Данные собираются непосредственно от пользователей по средствам GPS на смартфонах или других устройств, а такжес использованием видеокамер с высоким разрешением и дорожных радаров.

Использование подобных технологий позволяет получить информацию о скорости, расстоянии между транспортными средствами, маршрутах, движении через перекрестки, задержках и распределении между отдельными полосами движения и т.п. Затем данные попадают в центры управления дорожным движением, где происходит их анализ, и выстраиваются дальнейшие решения по оптимизации трафика. После этого обработанная информация отправляется на «умные» светофоры для смены сигналов в зависимости от загруженности соседних перекрестков, в случае ДТП на мобильные приложения для предупреждения о затруднениях проезда и координации общественного транспорта, а также др.

В качестве основных угроз, при оценке рисков функционирования ИСвыделим возможные векторы атак рассматриваемого фрагмента организации:

- анализ сетевого трафика;
- сканирование сети;
- угроза выявления пароля;
- подмена доверенного объекта сети;
- фишинг;
- внедрение ложного объекта сети;
- отказ в обслуживании;
- угрозы внедрения по сети вредоносных программ.

Исходя из сформированного списка векторов атак и последствий их реализации, вы-

Таблица 2.

Список концептов НКК

Концепт	Класс	Название концепта	Переменные состояния
C_1	Дестабилизирующие факторы (угрозы)	Угроза анализа сетевого трафика	Среднее количество реализации несанкционированного доступа, в ед. времени
C_2		Угроза сканирования сети	Среднее количество обнаруженных узлов, в ед. времени
C_3		Угроза выявления пароля	Среднее количество подобранных паролей, в ед. времени
C_4		Угроза подмены доверенного объекта сети	Среднее количество подмененных объектов, в ед. времени
C_5		Фишинг	Среднее количество фишинговых атак, в ед. времени
C_6		Угроза внедрения по сети вредоносных программ	Среднее количество вирусных атак, в ед. времени
C_7		Угроза физического выведения из строя средств хранения, обработки и (или) ввода/вывода/передачи информации	Среднее количество средств, выведенных из строя, в ед. времени
C_8		Угроза отказа в обслуживании	Среднее количество отказов, в ед. времени
C_9	Информационные ресурсы	АРМ сотрудников	Количество работоспособных АРМ, ед.
C_{10}		КПДК	Количество работоспособных КПДК, ед.
C_{11}		Программное обеспечение	Количество видов ПО, ед.
C_{12}		База данных	Объем информации, содержащейся в БД, Гбайт
C_{13}	Целевые факторы	Репутация организации	Число негативных высказываний, ед.
C_{14}		Качество предоставляемых услуг	Точность метеорологических показаний, %
C_{15}		Материальный ущерб	Финансовые потери, направленные на восстановление работоспособности, руб.

деляются концепты НКК карты ИС, актуальные для ИС «Метеорологическое обеспечение». Основные концепты когнитивной карты приведены в таблице 2.

Зададим веса W_{ij} с помощью нечеткой лингвистической шкалы, которая представляет собой упорядоченное множество лингвистических знаний (термов)[14]. Каждому из термов поставим в соответствие некоторый числовой диапазон, принадлежащий отрезку $[0;1]$ для обозначения положительного влияния или отрезку $[-1;0]$ - для отрицательных.

Таким образом, определяются входные данные для моделирования нечеткой когнитивной карты.

3. Построение графа с учетом особенностей ИС «Метеорологическое обеспечение»

В программной среде VUE построим модель оценки рисков НКК (рис. 4), в которой концепты обозначены номерами, а на связях между концептами определены значения весов. Значения весов определяются экспертным методом.

Рис. 4. Модель оценки рисков на основе НКК

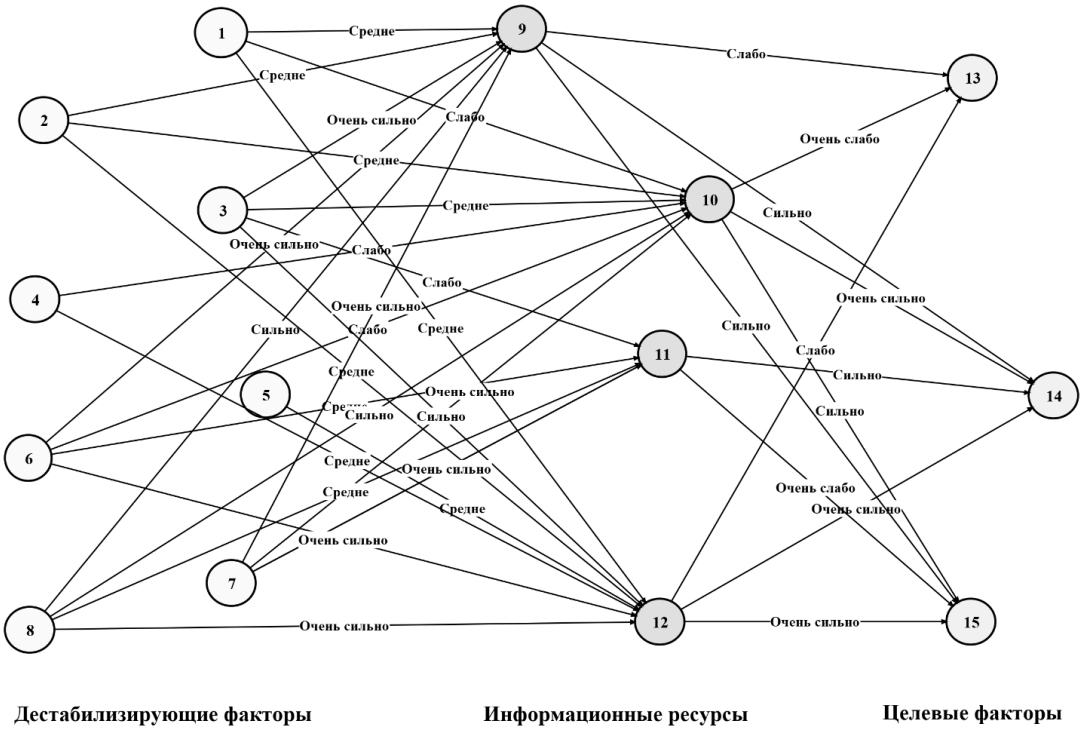

Рис. 4. Модель оценки рисков на основе НКК

Переведенныечисленные значения лингвистических термвесов связей НКК представлены в таблице 3.

Расчет состояния i-го концепта НКК производится уравнением вида[15]:

$$X_i(t+1) = f\left(\sum_{j=1}^{n} W_{ji}X_j(t) + X_i(t)\right) \qquad (1)$$

где $X_i(t+1)$ и $X_i(t)$ - значения переменной X_i, на $(t+1))$-м и t-м шаге соответственно, $(k = 1,2,...)$; W_{ji}- вес связи между концептами C_i и C_j; f - некоторая нелинейная функция, принимающая значения в интервале [0,1].

Таблица 3.

Веса связей между концептами НКК

Вес связи $C_i \rightarrow C_j$	НКК	Вес связи $C_i \rightarrow C_j$	НКК	Вес связи $C_i \rightarrow C_j$	НКК
W_{ij}	W_{ij}	$W_{4\rightarrow12}$	0,475	$W_{8\rightarrow12}$	0,925
$W_{1\rightarrow9}$	0,475	$W_{5\rightarrow12}$	0,475	$W_{9\rightarrow13}$	0,25
$W_{1\rightarrow10}$	0,25	$W_{6\rightarrow9}$	0,925	$W_{9\rightarrow14}$	0,725
$W_{1\rightarrow12}$	0,475	$W_{6\rightarrow10}$	0,25	$W_{9\rightarrow15}$	0,725
$W_{2\rightarrow9}$	0,475	$W_{6\rightarrow11}$	0,475	$W_{10\rightarrow13}$	0,1
$W_{2\rightarrow10}$	0,475	$W_{6\rightarrow12}$	0,925	$W_{10\rightarrow14}$	0,925
$W_{2\rightarrow12}$	0,475	$W_{7\rightarrow9}$	0,925	$W_{10\rightarrow15}$	0,725
$W_{3\rightarrow9}$	0,925	$W_{7\rightarrow10}$	0,925	$W_{11\rightarrow14}$	0,925
$W_{3\rightarrow10}$	0,925	$W_{7\rightarrow11}$	0,925	$W_{11\rightarrow15}$	0,1
$W_{3\rightarrow11}$	0,25	$W_{8\rightarrow9}$	0,725	$W_{12\rightarrow13}$	0,25
$W_{3\rightarrow12}$	0,725	$W_{8\rightarrow10}$	0,725	$W_{12\rightarrow14}$	0,925
$W_{4\rightarrow10}$	0,25	$W_{8\rightarrow11}$	0,475	$W_{12\rightarrow15}$	0,925

Таблица 4.

Результаты моделирования угроз

R_{13}	R_{14}	R_{15}
0,245	0,647	0,925

4. Расчет относительного уровня рисков для целевых факторов

Рассчитаем относительный уровень риска R_i для целевых факторов: репутация организации (C_{13}), качество предоставляемых услуг (C_{14}), материальный ущерб (C_{15}). Значение уровня угрозы рассчитывается как $R_i=X_i^*$, где X_i^* - установившееся значения состояния i-го целевого концепта, в данном случае i=13,14,15. Результаты моделирования уровня угроз целевых факторов показаны в таблице 4.

Для удобства интерпретации представим значения уровней угроз целевых концептов в виде диаграммы (рис. 5).

Полученные значения целевых факторов рисков информационной безопасности на целевые факторы позволяют выработать рекомендации по приоритетным направлениям для создания и/или совершенствования системы защиты информации. На основании диаграммы значений целевых факторов можно сделать вывод, что наиболее важным является необходимость проведения мероприятий по снижению риска, влекущего материальный ущерб организации. Следующим шагом по рейтингу опасности необходимо снизить уровень риска, влияющий на качество предоставляемых услуг. Уровень риска для репутации организации является приемлемым, поэтому, в данном случае, можно вынести решение о принятии риска.

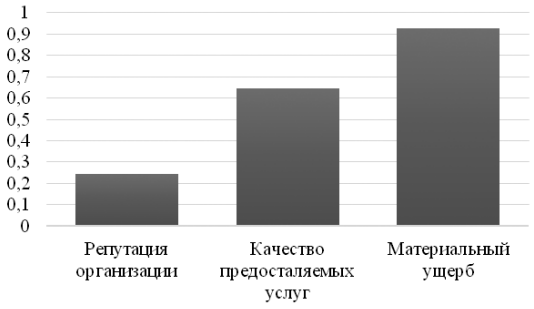

Рисунок 5 – Диаграмма значений целевых факторов

Заключение

Таким образом, приведенный в статье пример применения методики оценки рисков безопасности информации для критически важной информационной инфраструктуры ИС «Метеорологическое обеспечение» в организации ГКУ «ЦБДД» показывает, что оценка рисков, основанная на построении НКК– это эффективный инструмент для анализа нечеткости и неопределенности в оценке угроз безопасности информации и принятия обоснованных решений, что повышает эффективность системы безопасности, в целом. Кроме того, моделирование на основе НКК позволяет учитывать широкий спектр факторов и их взаимосвязей, а также использовать возможность визуализации оценки рисков информационной безопасности. Применение НКК может быть расширено, с учетом новых факторов, характеристик ИС и изменений предметной области. НКК предоставляют собой фундаментальную основу для дальнейшего развития и усовершенствования процесса оценки рисков информационной безопасности.

Литература

1. Шабуров, А.С. Разработка нечеткой когнитивной модели для автома-тизации оценки угроз безопасности информации / А.С. Шабуров, А.С. Ожгибесова // Master'sJournal. – 2023. –№ 2. – Art. № 05.

2. Ожгибесова А.С., Шабуров А.С.Метод автоматизации оценки угроз безопасности информации на основе нечетких когнитивных моделей // Иннова-ционные технологии: теория, инструменты, практика. – 2022. – Т.1. – С. 290-296.

3. Курчеева Г.И., Денисов В.В. Угрозы для информационной безопасности в высокоорганизованных системахтипа «Умный город» // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 8, №3 (2016) http://naukovedenie.ru/PDF/146EVN316.pdf (доступ свободный). (Дата обращения: 02.03.2024).

4. М.Р. Якимов. Транспортное планирование: терминологический сло-варь / М. Р. Якимов. – М: Агентство РАДАР, 2022. – 87 с. (Дата обращения: 29.02.2024).

5. Булдакова Т. И., Миков Д. А. Обеспечение согласованности и адек-ватности оценки факторов риска информационной безопасности // Вопросы ки-бербезопасности. 2017. №3 (21). URL: https://cyberleninka.ru/article/n/obespechenie-soglasovannosti-i-adekvatnosti-otsenki-faktorov-riska-informatsionnoy-bezopasnosti (дата обращения: 04.03.2024).

6. Васильев В.И. Интеллектуальные системы защиты информации: учебное пособие/. 2 е изд., испр. и доп. М.: Машиностроение, 2013. 172 с. (Дата обращения: 05.03.2024).

7. Васильев В.И., Вульфин А.М., Герасимова И.Б., Картак В.М. Анализ рисков кибербезопасности с помощью нечетких когнитивных карт.// Вопросы кибербезопасности. 2020. №2 (36). (дата обращения: 21.03.2024).

8. FCMapper // [Электронный ресурс] – Режим доступа: https://www.fcmappers.net/joomla/index.php (Дата обращения: 15.03.2024).

9. CmapTools // [Электронный ресурс] – Режим доступа: https://cmap.ihmc.us/ (Дата обращения: 16.03.2024).

10. XMind // [Электронный ресурс] – Режим доступа: https://xmind.app/ (Дата обращения: 17.03.2024).

11. VisualUnderstandingEnvironment// [Электронный ресурс] – Режим доступа: https://vue.tufts.edu/ (Дата обращения: 18.03.2024).

12. Федеральный закон от 07.07.2003 N 126-ФЗ (ред. 14.11.2023) "О связи". (Дата обращения: 09.03.2024).

13. Корнев, Л. В. Определение уровня безопасности системы защиты ин-формации на основе когнитивного моделирования / Л. В. Корнев. —— Текст: непосредственный // Молодой ученый. —— 2021. —— № 33 (375). —— С. 9--14. —— URL: https://moluch.ru/archive/375/83624/ (Дата обращения: 12.03.2024).

14. Васильев В. И., Вульфин А. М., Кудрявцева Р. Т. Анализ и управление рисками информационной безопасности с использованием технологий когнитивного моделирования // Доклады ТУСУР. 2017. №4. (Дата обращения: 19.03.2024).

15. Васильев В.И. Автоматизация процесса оценки информационных рисков с использованием нечетких когнитивных карт / В.И. Васильев, Р.Т., Кудрявцева, В.А. Юдинцев // Вестник УГАТУ, 2014. Т. 18, № 3 (64). С. 253 260. (дата обращения: 15.03.2024).

References

1. SHaburov, A.S. Razrabotkanechetkojkognitivnojmodelidlyaavto-matizaciiocenkiugrozbezopasnosti informacii / A.S. SHaburov, A.S. Ozhgibesova // Master'sJournal. – 2023. –№ 2. – Art. № 05.

2. Ozhgibesova A.S., SHaburov A.S. Metodavtomatizaciiocenkiugrozbezopasnostiinformaciinaosnove nechetkihkognitivnyhmodelej // Innovacionnyetekhnologii: teoriya, instrumenty, praktika. – 2022. – T.1. – S. 290-296.

3. Kurcheeva G.I., Denisov V.V. Ugrozydlyainformacionnojbezopas-nosti v vysokoorganizovannyhsiste mahtipa «Umnyjgorod» // Internet-zhurnal «NAUKOVEDENIE» Tom 8, №3 (2016) http://naukovedenie.ru/PDF/146EVN316.pdf (dostupsvobodnyj). (Data obrashche-niya: 02.03.2024).

4. M.R. YAkimov. Transportnoe planirovanie: terminologicheskij slo-var' / M. R. YAkimov. – M: Agentstvo RADAR, 2022. – 87 s. (Data obrashcheniya: 29.02.2024).

5. Buldakova T. I., Mikov D. A. Obespecheniesoglasovannosti i adek-vatnostiocenkifaktorovriskainformacionnojbezopasnosti // Voprosyki-berbezopasnosti. 2017. №3 (21). URL: https://cyberleninka.ru/article/n/obespechenie-soglasovannosti-i-adekvatnosti-otsenki-faktorov-riska-informatsionnoy-bezopasnosti (data obrashcheniya: 04.03.2024).

6. Vasil'ev V.I. Intellektual'nye sistemy zashchity informacii: uchebnoe posobie/. 2 e izd., ispr. i dop. M.: Mashinostroenie, 2013. 172 s. (Data obrashcheniya: 05.03.2024).

7. Vasil'ev V.I., Vul'fin A.M., Gerasimova I.B., Kartak V.M. Ana-liz riskov kiberbezopasnosti s pomoshch'yu nechetkih kognitivnyh kart. // Vo-prosy kiberbezopasnosti. 2020. №2 (36). (data obrashcheniya: 21.03.2024).

8. FCMapper // [Elektronnyj resurs] – Rezhim dostupa: https://www.fcmappers.net/joomla/index.php (Data obrashcheniya: 15.03.2024).

9. CmapTools // [Elektronnyj resurs] – Rezhim dostupa: https://cmap.ihmc.us/ (Data obrashcheniya: 16.03.2024).

10. XMind // [Elektronnyj resurs] – Rezhim dostupa: https://xmind.app/ (Data obrashcheniya: 17.03.2024).

11. Visual Understanding Environment // [Elektronnyj resurs] – Rezhim dostupa: https://vue.tufts.edu/ (Data obrashcheniya: 18.03.2024).

12. Federal'nyj zakon ot 07.07.2003 N 126-FZ (red. 14.11.2023) "O svyazi". (Data obrashcheniya: 09.03.2024).

13. Kornev, L. V. Opredelenie urovnya bezopasnosti sistemy zashchity informacii na osnove kognitivnogo modelirovaniya / L. V. Kornev. —— Tekst: neposredstvennyj // Molodoj uchenyj. —— 2021. —— № 33 (375). —— S. 9--14. —— URL: https://moluch.ru/archive/375/83624/ (Data obrashcheniya: 12.03.2024).

14. Vasil'ev V. I., Vul'fin A. M., Kudryavceva R. T. Analiz i uprav-lenie riskami informacionnoj bezopasnosti s ispol'zovaniem tekhnologij ko-gnitivnogo modelirovaniya // Doklady TUSUR. 2017. №4. (Data obrashcheniya: 19.03.2024).

15. Vasil'ev V.I. Avtomatizaciyaprocessaocenkiinformacionnyhriskov s ispol'zovaniemnechetkihkognit ivnyhkart / V.I. Vasil'ev, R.T., Kudryavceva, V.A. YUdincev // Vestnik UGATU, 2014. T. 18, № 3 (64). S. 253 260. (dataobrashcheniya: 15.03.2024).

ОЖГИБЕСОВА Анна Сергеевна, аспирант кафедры Автоматики и теле-механики, Пермский национальный исследовательский политехнический универ-ситет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: aozgibesova@pstu.ru

ШАБУРОВ Андрей Сергеевич, кандидат технических наук, доцент, доцент кафедры Автоматики и телемеханики, Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: shans@at.pstu.ru

ЮЖАКОВ Александр Анатольевич, доктор технических наук, профессор, заведующий кафедры Автоматики и телемеханики, Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: uz@at.pstu.ru

OZHGIBESOVA Anna Sergeevna, graduatestudent, Department of Automation and Telemechanics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: aozgibesova@pstu.ru

SHABUROV Andrey Sergeevich, Candidate of Technical Sciences, Associate Professor of the Department of Automation and Telemechanics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: shans@at.pstu.ru

YUZHAKOV Alexander Anatolyevich, Doctor of Technical Sciences, Profes-sor, Head of the Department of Automation and Telemechanics, Perm National Re-search Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: uz@at.pstu.ru

УДК 007.51 Вестник УрФО № 2(52) / 2024, с. 68–76

Быкасов А. В., Соколов А. Н. DOI: 10.14529/secur240207

СТАТИСТИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ ВЛИЯНИЯ КОМБИНИРОВАННОГО МЕТОДА АКТИВНОГО СКАНИРОВАНИЯ НА СТАБИЛЬНОСТЬ ФУНКЦИОНИРОВАНИЯ СЕТИ АСУ ТП

Методы активного сканирования все шире применяются в системах обнаружения вторжений (СОВ) для сбора и анализа сетевой информации. Поскольку активное сканирование предполагает отклик от устройств сети, это может повлиять на стабильность работы сетевого оборудования, что является критичным для промышленных сетей.

В работе представлена статистическая модель оценки влияния комбинированного метода активного сканирования на стабильность функционирования сети АСУ ТП. Проведены эксперименты для оценки влияния комбинированного метода активного сканирования на стабильность функционировании сети АСУ ТП. Полученные результаты, представленные в сравнении с результатами для распространенного инструмента сетевого сканирования Nmap, дают сделать вывод о том, что комбинированный метод активного сканирования оказывает меньшее влияние на сеть, а также обеспечивает стабильность при сканировании, тем самым не нарушая работоспособность сети АСУ ТП.

Ключевые слова: *автоматизированная система управления технологическим процессом (АСУ ТП), активное сканирование, система обнаружения вторжений (СОВ), контроль сетевых устройств, протокол определения адреса (ARP), протокол межсетевых управляющих сообщений (ICMP).*

STATISTICAL MODEL FOR ASSESSING THE IMPACT OF THE COMBINED ACTIVE SCANNING METHOD ON THE STABILITY OF THE INDUSTRIAL CONTROL NETWORK

Active scanning methods are increasingly used in intrusion detection systems (IDS) to collect and analyze network information. Since active scanning involves response from network devices, this can affect the stability of network equipment, which is critical for industrial networks.

The paper presents a statistical model for assessing the influence of the combined active scanning method on the stability of the functioning of the industrial control network. Experiments were carried out to evaluate the influence of the combined active scanning method on the stability of the industrial control network. The results obtained, presented in comparison with the results for the common network scanning tool Nmap, lead to the conclusion that the combined method of active scanning has less impact on the network, and also provides stability during scanning, thereby not disrupting the functionality of the ICS network.

Keywords: *industrial control system (ICS), active scanning, intrusion detection system (IDS), control of network devices, address resolution protocol (ARP), internet control message protocol (ICMP).*

Для обеспечения информационной безопасности сетей автоматизированных систем управления технологическими процессами (АСУ ТП) применяются системы обнаружения и предотвращения вторжений (СОВ и СПВ). Эти системы предназначены для мониторинга целостности сетевого трафика и программного обеспечения, функционирующего на узлах сети. Данный функционал может быть реализован полностью посредством пассивного анализа сети, включающего сниффинг и пассивный анализ трафика. Однако такой метод требует значительных временных затрат и в некоторых случаях может не обеспечивать достаточную точность. Поэтому системы обнаружения вторжений часто дополняются методами активного сканирования [1], суть которых заключается в отправке СОВ специально сформированных запросов целевым узлам, всем узлам, периферийному оборудованию или всем типам устройств, а затем интерпретации полученных ответов [2].

Более оперативное получение информации об используемых устройствах, установленном ПО и конфигурациях в сети АСУ ТП обеспечивается путем интеграции СОВ с установленным на рабочие станции специализированным программным обеспечением — агентом инвентаризации, который передает информацию либо непосредственно в СОВ, либо через систему управления информационной безопасностью (SIEM). Этот метод применим только к устройствам, на которые возможно свободно устанавливать прикладное ПО. Однако множество встраиваемых систем в АСУ ТП, таких как программируемые логические контроллеры (ПЛК), распределённый ввод-вывод, сетевое оборудование и т.д., обычно не позволяют установить дополнительное ПО [3 – 5].

Таким образом, возникает необходимость применения активного сканирования, запускаемого средствами самой СОВ или сторонними ресурсами. Однако такие методы

нарушают принцип невмешательства СОВ в рабочий процесс сети АСУ ТП для получения информации. Следовательно, для обеспечения информационной безопасности сетей АСУ ТП и одновременной минимизации вмешательства в их рабочий процесс необходима разработка специализированных методов активного сканирования [6,7].

Целью работы является разработка статистической модели для оценки влияния методов активного сканирования на стабильность функционирования сети АСУ ТП. Модель должна позволять на основе статистических данных работы метода активного сканирования делать заключение о пригодности метода для его применения в сетях АСУ ТП.

Для оценки метода берутся такие показатели, как время выполнения каждого этапа сканирования (если это возможно), среднее время сканирования, а также средний показатель всплесков трафика, возникающих при сканировании. В качестве статистических показателей для оценки работы модели рассчитываются:

выборочное математическое ожидание

$$M = \frac{\sum_{i=1}^{n} a[i]}{n} \quad (1)$$

где $a[i]$ – элемент экспериментально полученного статистического набора данных, n – объем выборки;

и выборочная дисперсия:

$$D = \frac{\sum_{i=1}^{n} (a[i]-M)^2}{n-1} \quad (2)$$

где $a[i]$ – элемент выборки, n – объем выборки.

Рассмотренный комбинированный метод активного сканирования [8] состоит из 4-х этапов сканирования:

1. ARP-сканирование. На этом этапе отправляется широковещательный запрос в сеть с целью обнаружения подключенных хостов, а также нахождения их IP- и MAC-адресов. В качестве входных данных используется подсеть или IP-адрес устройства. В результате получается список устройств, содержащий IP- и MAC-адреса.

2. CMP-сканирование. В качестве входных данных используется список IP-адресов, полученный на предыдущем этапе. Далее, на каждое устройство из списка отправляется ping-запрос. Если ответ приходит, мы считаем это устройство «живым». В результате получаем список «живых» IP-адресов.

3. SNMP-сканирование. На данном этапе проводится опрос устройств по протоколу SNMP с целью получения информации об устройстве: марка, модель, имя в сети и т.д. В качестве входных данных используется результат ICMP-сканирования.

4. Сканирование портов. На данном этапе проводится поиск открытых портов устройств посредством отправки TCP-пакетов. Сканирование может проводится как по определенному заранее списку «популярных» портов, так и по заданному вручную. В качестве входных данных также используется список IP-адресов, полученный на этапе ICMP-сканирования.

Nmap – инструмент для сканирования сетей, поддерживающий как пассивные, так и активные методы активного сканирования. Он может быть использован для проверки

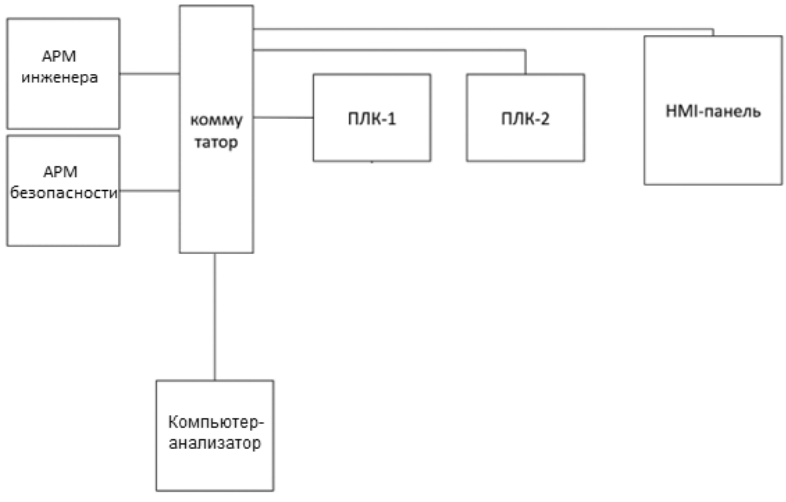

Рис. 1. Схема сетевых соединений первого сегмента лабораторного стенда

Рис. 2. Схема сетевых соединений второго сегмента лабораторного стенда

безопасности, просто для определения сервисов, запущенных на узле, для идентификации ОС и приложений и т.д.

Экспериментальные исследования проводились на лабораторном стенде, состоящем из двух сегментов. Первый сегмент осуществляет сетевое взаимодействие двух ПЛК (рис. 1). Состав оборудования:

- ПЛК-1, Siemens-1512;
- ПЛК-2, Siemens-1510;
- Коммутатор Scalance XC208;
- АРМ инженера, содержащая ПО для программирования ПЛК;
- АРМ оператора системы безопасности;
- HMI- панель для визуализации и управления.
- На ПЛК загружена программа, эмулирующая металлонарезной станок.

Для проверки методов сканирования к стенду подключен ноутбук, исполняющий роль внешнего сканирующего устройства.

Второй сегмент лабораторного стенда состоит из одного ПЛК, на котором эмулируется программа управления насосами. Состав оборудования второго сегмента сети (рис. 2):

- ПЛК KEAZ OptiLogic L-CPU-2-M;
- Коммутатор;
- АРМ инженера, содержащая ПО для программирования ПЛК;
- АРМ оператора системы безопасности;
- HMI- панель для визуализации и управления.

Для статистической оценки влияния метода активного сканирования на промышленную сеть проведен ряд экспериментов для накопления статистической базы. Для каждого сегмента лабораторного стенда проведены сканирования с использованием комбинированного метода активного сканирования и инструмента Nmap.

При сканировании первого сегмента сети оба инструмента показали следующий результат: все хосты в сети были найдены, все ответили на ping-запрос, хостам присвоен статус «жив». От двух хостов получена информация путем сканирования методом SNMP. Все открытые порты были найдены. При этом, рассмотренные методы сканирования оказали различное влияние на сеть.

На рисунке 3 представлены статистические результаты для среднего показания всплесков трафика, возникавших при сканированиях. Верхняя кривая соответствует результатам сканирования с использованием Nmap (выборочное математическое ожидание равно 1835,7, выборочная дисперсия – 84427,41). Нижняя кривая соответствует результатам сканирования с использованием комбинированного метода активного сканирования (выборочное математическое ожидание равно 184,733, выборочная дисперсия – 116,929). Из графиков видно, что выборочное математическое ожидание среднего показателя всплесков для Nmap на порядок больше, чем для комбинированного метода. Поэтому можно сделать заключение о том, что Nmap оказывает значительно большее влияние на сеть, чем комбинированный метод. Выборочная дисперсия для комбинированного метода в несколько раз меньше, чем у Nmap. Следовательно, комбинированный метод показывает более стабильные результаты.

На рисунке 4 представлены статистические результаты общего времени сканирования. Верхняя кривая соответствует результа-

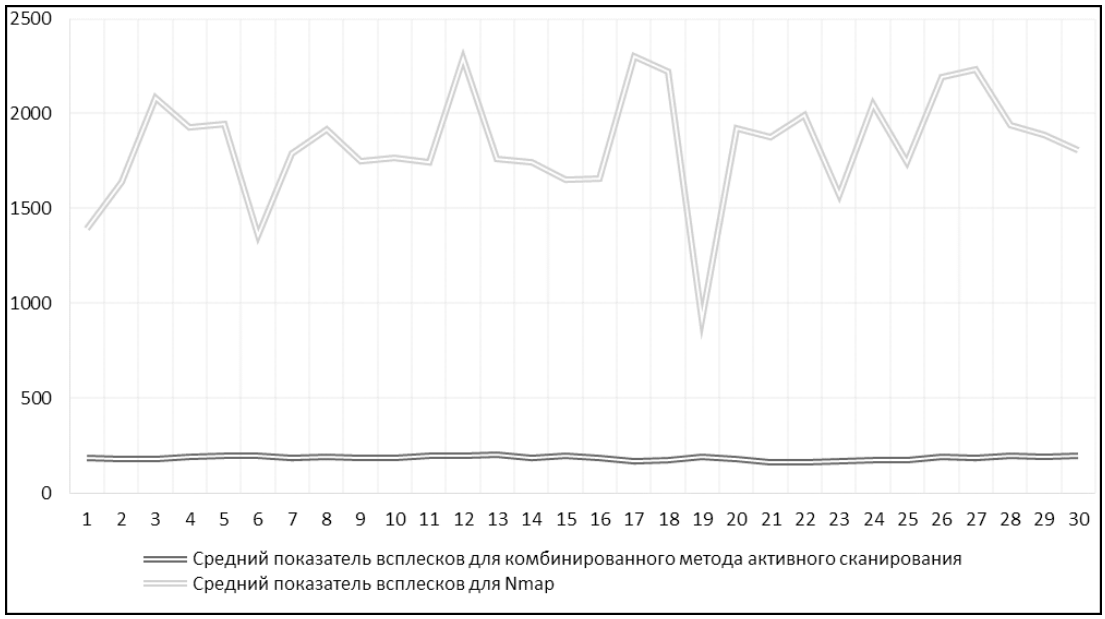

Рис. 3. Сравнение среднего показателя всплесков для 1 сегмента сети

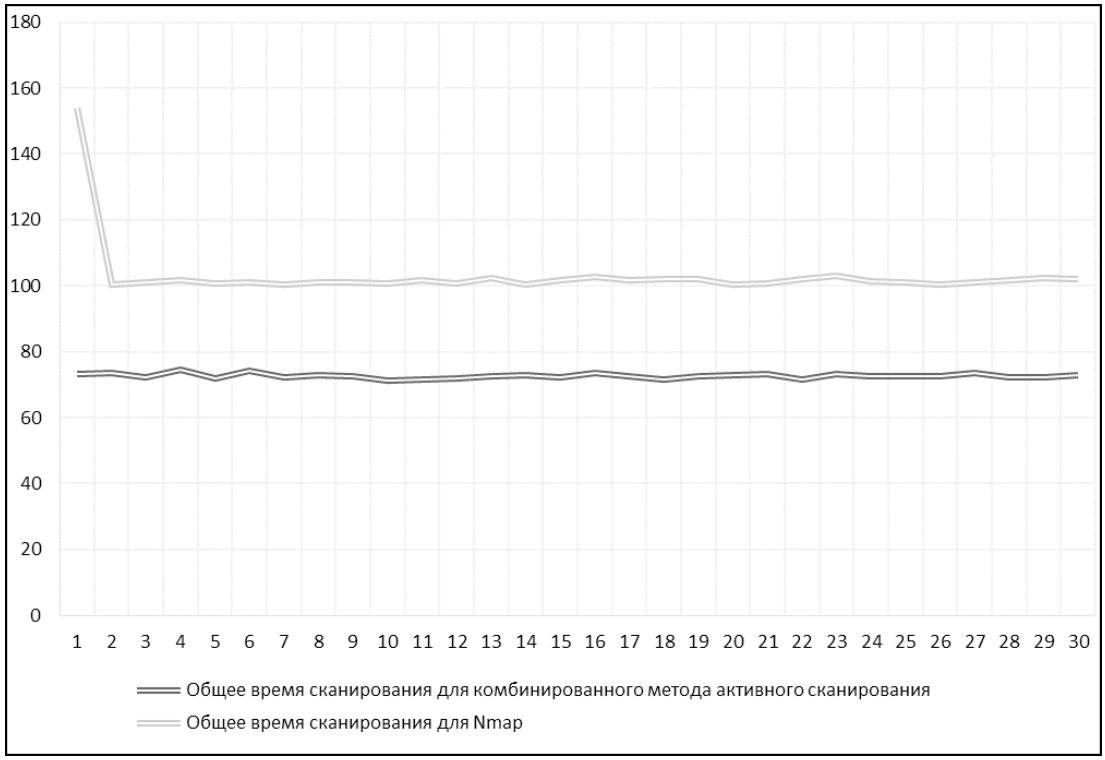

Рис. 4. Сравнение общего времени сканирования для 1 сегмента сети

там сканирования с использованием Nmap (выборочное математическое ожидание равно 103,129, выборочная дисперсия – 89,593). Нижняя кривая соответствует результатам сканирования с использованием комбинированного метода активного сканирования (выборочное математическое ожидание равно 72,714, выборочная дисперсия – 0,596.). Из графиков видно, что выборочное математическое ожидание общего времени сканирования для Nmap больше, чем для комбинированного метода. Выборочная дисперсия для комбинированного метода в несколько раз меньше, чем у метода Nmap. Следовательно, комбинированный метод показывает более стабильные результаты сканирования, и про-

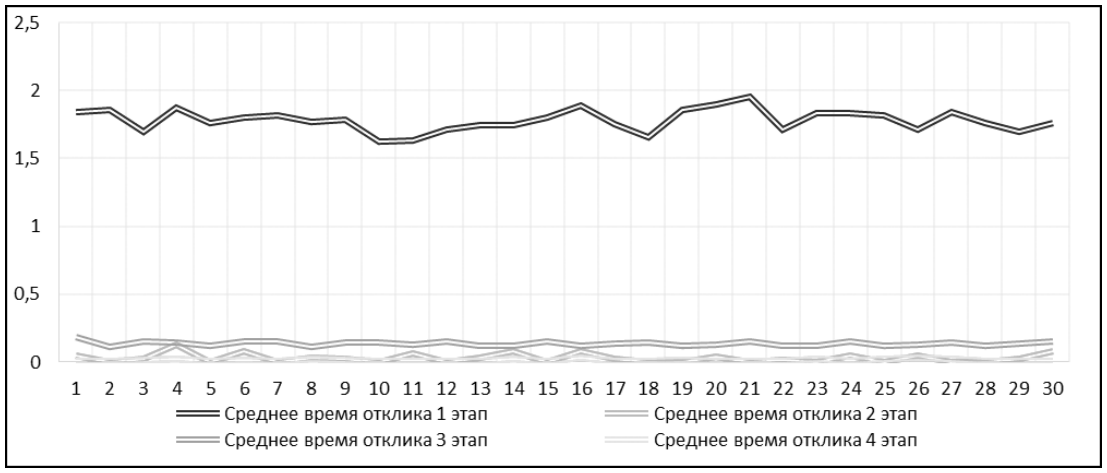

Рис. 5. Сравнение среднего времени отклика для этапов комбинированного метода активного сканирования для 1 сегмента сети

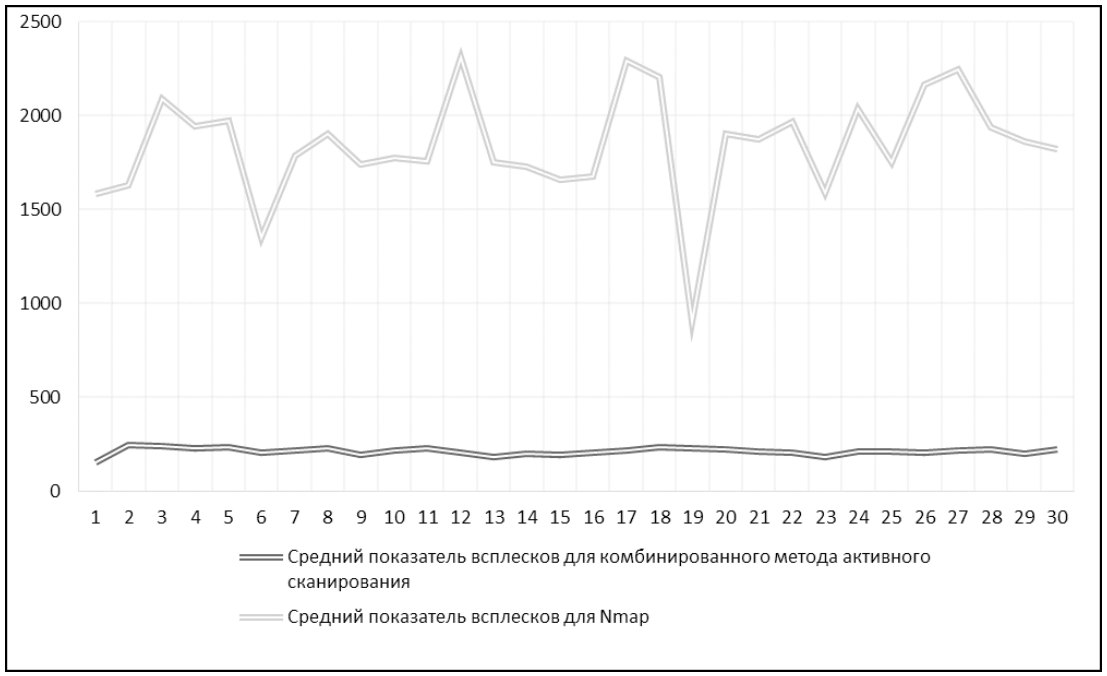

Рис. 6. Сравнение среднего показателя всплесков для 2 сегмента сети

должительность сканирования при этом меньше, чем у Nmap.

На рисунке 5 представлены статистические результаты среднего времени отклика для этапов комбинированного метода активного сканирования. Их выборочные математические ожидания небольшие, а выборочные дисперсии близки к нулю. Следовательно, можно сделать заключение о предсказуемости результатов работы комбинированного метода с точки зрения минимального влияния на работу сети. При этом Nmap не позволяет оценить среднее время отклика.

При сканировании второго сегмента сети оба инструмента показали следующий результат: все хосты в сети были найдены, все ответили на ping-запрос, хостам присвоен статус «жив». Поскольку хосты не настроены на взаимодействие по протоколу SNMP, получить информацию о них невозможно. Все открытые порты были найдены. При этом рассмотренные методы сканирования оказали различное влияние на сеть.

На рисунке 6 представлены статистические результаты для среднего показания всплесков трафика, возникавших при сканированиях. Верхняя кривая соответствует результатам сканирования с использованием Nmap (выборочное математическое ожидание равно 1840,693, выборочная дисперсия - 80926,009). Нижняя

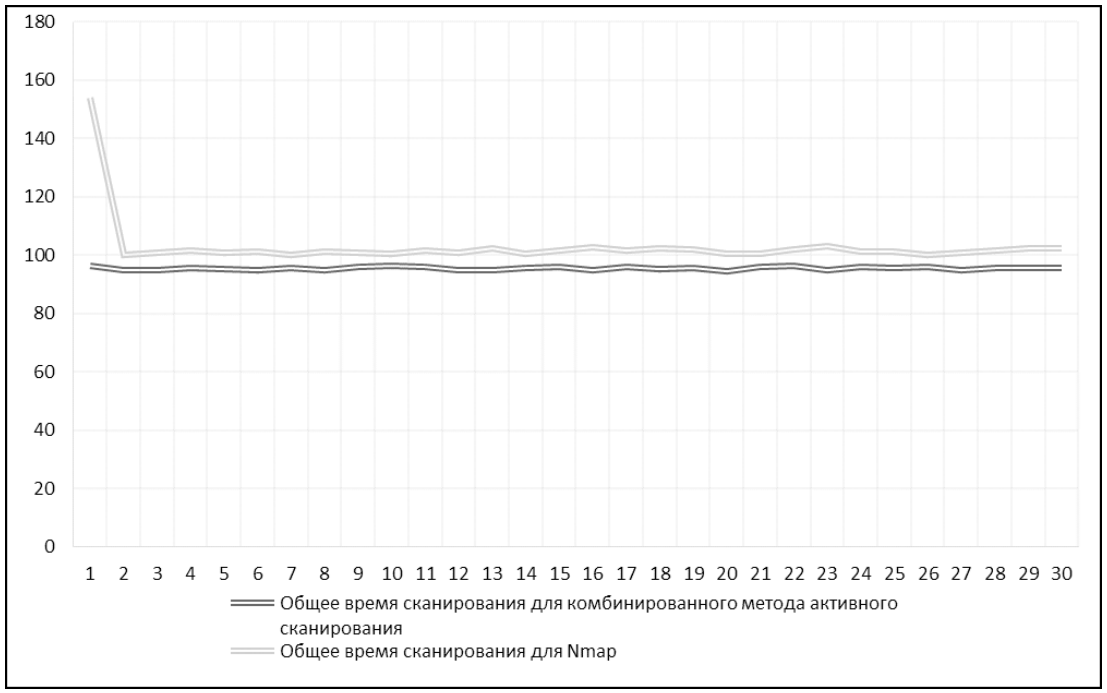

Рис. 7. Сравнение общего времени сканирования для 2 сегмента сети

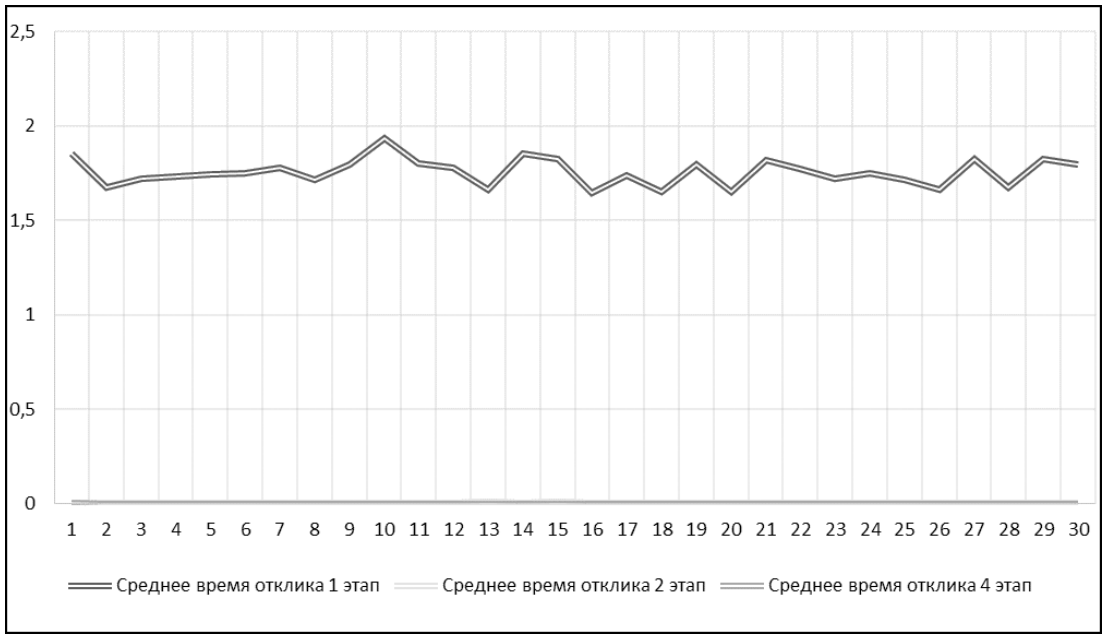

Рис. 8. Сравнение среднего времени отклика для этапов комбинированного метода активного сканирования для 2 сегмента сети

кривая соответствует результатам сканирования с использованием комбинированного метода активного сканирования (выборочное математическое ожидание равно 212,7, выборочная дисперсия - 358,477). Из графиков видно, что выборочное математическое ожидание среднего показателя всплесков для Nmap на порядок больше, чем для комбинированного метода. Поэтому можно сделать заключение о том, что Nmap оказывает значительно большее влияние на сеть, чем комбинированный метод. Выборочная дисперсия для комбинированного метода в несколько раз меньше, чем у Nmap. Следовательно, комбинированный метод показывает более стабильные результаты.

На рисунке 7 представлены статистические результаты общего времени сканирования. Верхняя кривая соответствует результатам сканирования с использованием Nmap (выборочное математическое ожидание равно 103,348, выборочная дисперсия - 84,847). Нижняя кривая соответствует результатам сканирования с использованием комбинированного метода активного сканирования (выборочное математическое ожидание равно 95,513, выборочная дисперсия - 0,295). Из графиков видно, что выборочное математическое ожидание общего времени сканирования для Nmap больше, чем для комбинированного метода. Выборочная дисперсия для комбинированного метода в разы меньше, чем у Nmap. Следовательно, комбинированный метод показывает более стабильные результаты сканирования, и продолжительность сканирования при этом меньше, чем у Nmap.

На рисунке 8 представлены статистические результаты среднего времени отклика для этапов комбинированного метода активного сканирования. Их выборочные математические ожидания небольшие, а выборочные дисперсии близки к нулю. Следовательно, можно сделать заключение о предсказуемости результатов работы комбинированного метода с точки зрения минимального влияния на работу сети. При этом Nmap не позволяет оценить среднее время отклика.

Исходя из результатов экспериментов можно сделать вывод, что комбинированный метод активного сканирования имеет следующие статистические характеристики, в сравнении с Nmap. Выборочное математическое ожидание для среднего показателя всплесков и для общего времени сканирования меньше, чем у Nmap. Следовательно, комбинированный метод активного сканирования оказывает меньше влияния на сеть, а также занимает меньше времени при сканировании сети. Выборочная дисперсия для среднего показателя всплесков и для общего времени сканирования значительно меньше, чем у Nmap. Из этого можно сделать вывод, что комбинированный метод активного сканирования показывает стабильные результаты сканирования, тем самым не нарушая работоспособность сети АСУ ТП.

Литература

1. Павленко А. Сканирование на наличие уязвимостей. / А. Павленко // Отус онлайн-образование. – 2022. – URL: https://otus.ru/nest/post/2468/ (дата обращения 21 апреля 2024 г.)

2. Проведение активных опросов устройств с помощью Kaspersky Industrial CyberSecurity for Networks. - https://support.kaspersky.com/ KICSforNetworks/4.0/ru-RU/236044.htm (дата обращения 21 апреля 2024 г.)

3. Активное сканирование с помощью Nozomi Networks Guardian. - https://www.nozominetworks.com/downloads/US/Nozomi-Networks-Smart-Polling-Data-Sheet.pdf (дата обращения 21 апреля 2024 г.)

4. Hansson A. Analyzing Internet-connected industrial equipment. / A. Hansson, M. Khodari, A. Gurtov. // 2018 International Conference on Signals and Systems (ICSigSys). – 2018. – С. 29-35. – DOI: 10.1109/ICSIGSYS.2018.8372775 – URL: https://www.researchgate.net/publication/325635836_Analyzing_Internet-connected_industrial_equipment (дата обращения 21 апреля 2024 г.)

5. Исследование: более 4 000 устройств АСУ ТП уязвимы для удаленных атак / InfoWatch. – 2021. – URL: https://www.infowatch.ru/resources/blog /issledovanie-bolee-4-000-ustroystv-asu-tp-uyazvimy-dlya-udalennykh-atak (дата обращения 21 апреля 2024 г.)

6. Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью. Часть 1. Общие требования [Текст]: ГОСТ Р МЭК 61508-1-2012. – Введ.2013-08-01. – М.: Федеральное агентство по техническому регулированию и метрологии, 2012. – 586 с.

7. Функциональная безопасность систем электрических, электронных, программируемых электронных, связанных с безопасностью. Часть 3. Требования к программному обеспечению [Текст]: ГОСТ Р МЭК 61508-3-2012. – Введ. 2013-08-01. – М.: Федеральное агентство по техническому регулированию и метрологии, 2012. – 588 с.

8. Обеспечение безопасности сети АСУ ТП при использовании комбинированного метода активного сканирования // БЕЗОПАСНОСТЬ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА. Сборник научных трудов XXII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. 2024. – С. 159-166.

References

1. Pavlenko A. Skanirovaniye na nalichiye uyazvimostey. / A. Pavlenko // Otus onlayn-obrazovaniye. – 2022. – URL: https://otus.ru/nest/post/2468/ (data obrashcheniya 21 aprelya 2024 g.)

2. Provedeniye aktivnykh oprosov ustroystv s pomoshch'yu Kaspersky Industrial CyberSecurity for Networks. Rezhim dostupa: https://support.kaspersky.com /KICSforNetworks/4.0/ru-RU/236044.htm (data obrashcheniya 21 aprelya 2024 g.)

3. Aktivnoye skanirovaniye s pomoshch'yu Nozomi Networks Guardian. Rezhim dostupa: https://www.nozominetworks.com/downloads/US/Nozomi-Networks-Smart-Polling-Data-Sheet.pdf (data obrashcheniya 21 aprelya 2024 g.)

4. Hansson A. Analyzing Internet-connected industrial equipment. / A. Hansson, M. Khodari, A. Gurtov. // 2018 International Conference on Signals and Systems (ICSigSys). – 2018. – C. 29-35. – DOI: 10.1109/ICSIGSYS.2018.8372775 – URL: https://www.researchgate.net/publication/325635836_Analyzing_Internet-connected_industrial_equipment (data obrashcheniya 21 aprelya 2024 g.)

5. Issledovaniye: boleye 4 000 ustroystv ASU TP uyazvimy dlya udalennykh atak / InfoWatch. – 2021. – URL: https://www.infowatch.ru/resources/blog /issledovanie-bolee-4-000-ustroystv-asu-tp-uyazvimy-dlya-udalennykh-atak (data obrashcheniya 21 aprelya 2024 g.)

6. Funktsional'naya bezopasnost' sistem elektricheskikh, elektronnykh, programmiruyemykh elektronnykh, svyazannykh s bezopasnost'yu. Chast' 1. Obshchiye trebovaniya [Tekst]: GOST R MEK 61508-1-2012. – Vved. 2013-08-01. – M.: Federal'noye agentstvo po tekhnicheskomu regulirovaniyu i metrologii, 2012. – 586 s.

7. Funktsional'naya bezopasnost' sistem elektricheskikh, elektronnykh, programmiruyemykh elektronnykh, svyazannykh s bezopasnost'yu. Chast' 3. Trebovaniya k programmnomu obespecheniyu [Tekst]: GOST R MEK 61508-3-2012. – Vved. 2013-08-01. – M.: Federal'noye agentstvo po tekhnicheskomu regulirovaniyu i metrologii, 2012. – 588 s.

8. Obespecheniye bezopasnosti seti ASU TP pri ispol'zovanii kombinirovannogo metoda aktivnogo skanirovaniya // BEZOPASNOST' INFORMATSIONNOGO PROSTRANSTVA. Sbornik nauchnykh trudov XXII Vserosciyskoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh. 2024. – S. 159-166.

СОКОЛОВ Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой «Защита информации» федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, Уральский федеральный округ, Челябинская область, г. Челябинск, просп. В.И. Ленина, д. 76. E-mail: sokolovan@susu.ru

БЫКАСОВ Андрей Витальевич, аспирант федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, Уральский федеральный округ, Челябинская область, г. Челябинск, просп. В.И. Ленина, д. 76. E-mail: andreybikasov@gmail.com

SOKOLOV Alexander Nikolaevich, Candidate of Technical Sciences, Associate Professor, Head of the Information Security Department of the Federal State Autonomous Educational Institution of Higher Education South Ural State University (National Research University). 454080, Ural Federal District, Chelyabinsk Region, Chelyabinsk, prosp. IN AND. Lenina, d. 76. E-mail: sokolovan@susu.ru

BYKASOV Andrey Vitalievich, post-graduate student of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (National Research University)". 454080, Ural Federal District, Chelyabinsk Region, Chelyabinsk, prosp. IN AND. Lenina, d. 76. E-mail: andreybikasov@gmail.com

УДК 004.056.53 Вестник УрФО № 2(52) / 2024, с. 77–88

Басалай К. А., Зулькарнеев И. Р. DOI: 10.14529/secur240208

ПОДХОДЫ К КЛАССИФИКАЦИИ ВЕКТОРОВ АТАК И УЯЗВИМОСТЕЙ JSON WEB TOKENS[1]

В статье рассмотрены вопросы обеспечения безопасности JSON Web Tokens при их использовании в микросервисной архитектуре, мобильных и веб-приложениях. Обозначены преимущества использования JWT по сравнению с другими методами аутентификации и авторизации. Авторами поднята проблема реализации системного и комплексного подхода по обеспечению безопасности при использовании JWT и отсутствия какой-либо классификации уязвимостей и векторов атак на JWT в отечественной и зарубежной научно-технической литературе.

В ходе работы был исследован полный жизненный цикл JWT и их структура, используемые стандартные поля и алгоритмы формирования подписи, была составлена классификация уязвимостей на основе этапов жизненного цикла токена, и для каждого класса уязвимостей определены вектора атак. Были проанализированы существующие 119 уязвимостей базы CVE, связанных с JWT, определены новые категории и построены корреляции по данным с 2015г. На основе этого анализа авторами была предложена классификация векторов атак на основе жизненного цикла JWT и объектов атаки, а также выявлены новые уязвимости, которые не рассмотрены в научных работах на данный момент.

Данная классификация может быть использована при проектировании информационных систем на базе микросервисной архитектуры для обеспечения безопасности процессов аутентификации и авторизации на стороне клиента и сервера, определения слабых мест работающих сервисов, а также подбора инструментария для тестирования безопасности JWT.

Ключевые слова: *JWT, CVE, классификация уязвимостей, вектор атаки, аутентификация, авторизация, микросервисы.*

[1] Работа выполнена в рамках гранта РФФИ 20-47-720005

APPROACHES TO CLASSIFYING ATTACK VECTORS AND VULNERA-BILITIES OF JSON WEB TOKENS

Security issues of JSON web tokens (JWT) usage in a microservice architec-ture, in mobile and web applications are presented in the paper. It also outlines the benefits of JWT using over other authentication and authorization methods. The au-thors raise the problem of implementing a systematic and comprehensive approach to security assurance while using JWT. They also highlight the lack of national and international scientific and technical resources classifying JWT vulnerabilities and attack vectors.

The entire lifecycle and structural components of JSON Web Tokens, as well as the standard claims and algorithms used for signature generation were investigat-ed. In addition, the paper presents a classification of vulnerabilities based on JWT lifecycle stages, and defines attack vectors for each vulnerability class. The authors analyzed all existing 119 CVE vulnerabilities associated with JWTs, identified new categories and built correlations using data since 2015. Through the analysis the au-thors proposed the attack vectors classification based on the JWT lifecycle and attack targets, and identified new vulnerabilities that have not been considered in previous researches.

There are several ways to use the developed classification: to ensure the security of authentication and authorization processes on the client and server sides in the information systems design based on microservice architecture; to find vulnerabilities in the services; to select tools for testing the security of JWT.

Keywords: *JWT, CVE, vulnerability classification, attack vector, authentica-tion, authorization, microservices.*

Введение

Мобильные или веб-приложения являются неотъемлемой частью функционирования любой организации, стартапа или государственного учреждения. Крупные организации с большим количеством внутренних и внешних пользователей стараются создавать свои собственные информационные экосистемы, используя в том числе веб-сервисы и внутренние вычислительные ресурсы. Согласно исследованию [1] большинство организаций используют для этих целей микросервисную архитектуру, отмечая ее гибкость, масштабируемость, быстроту внесения изменений и удобство для разработчиков. Необходимость доступа к постоянно растущему числу сервисов большого количества пользователей поставило задачу использования быстрой, удобной и ресурсоемкой системы аутентификации, в том числе сквозной для всей экосистемы. Одним из самых распространенных решений данной проблемы является использование JSON Web Tokens (далее – JWT) [2]. Спектр применения этих токенов настолько высок, что они используются даже для управления доступом и обеспечения безопасности устройств IoT [3], что может быть применено при построении и управлении доступом, например, в медицинских информационных системах, клинико-диагностических лабораториях и телемедицине.

Популярность использования JWT обусловлена следующими преимуществами перед остальными методами аутентификации и авторизации:

Удобство – использование токенов упрощает доступ пользователей к большому количеству сервисов за счет сквозной аутентификации [4].

Автоматическая проверка пользователя – при каждом запросе от поль-зователя на сервер, его JWT помещается в заголовок http-пакета, что позволяет автоматически пройти аутентификацию [5].

Повышение скорости обработки запросов – JWT содержит в себе все необходимые пользовательские данные и передает их в открытом виде, снижая нагрузку на сервер [2, 5].

Гибкость – разработчики могут создавать собственные поля внутри токена, самостоятельно определять их содержимое и вариант применения [2].

Наряду с очевидными преимуществами JWT остро встает вопрос об обеспечении безопасности при их использовании, так как особенности структуры и методов реализации JSON Web Tokens служат причиной большого количества потенциальных уязвимостей, что приводит к появлению новых векторов атак на систему аутентификации и ее компрометации.

Целью данной статьи является проведение комплексного анализа существующих проблем JWT и составление классификации уязвимостей и векторов атак на основе их жизненного цикла. Классификация позволит сформировать целостное представление о возможных атаках на протяжении всех этапов использования JWT, что позволит обеспечить безопасность процессов аутентификации и авторизации; рассмотреть слабые места работающих сервисов, а также подобрать инструментарий для тестирования безопасности JWT.

Обзор предыдущих работ

Коллектив авторов статьи [5] описывает механизмы аутентификации и авторизации на основе JWT, рассматривает состав токена, выделяет базовые уязвимости JWT и методы их устранения, а также отмечает преимущества и недостатки использования JWT. Статья носит преимущественно обзорный характер и предложенные авторами способы обеспечения безопасности JWT охватывают лишь малую часть известных уязвимостей.

Авторы научной работы [4] сравнили механизмы аутентификации и авторизации на основе JWT и на основе сессий. Для обеспечения безопасности системы аутентификации авторы предлагают реализовать защиту от XSS-атак на стороне клиента, а также применение сложных механизмов валидации и использование сложных ключей шифрования на стороне сервера. Указанные ме-ры носят точечный характер и охватывают только некоторые атаки на JWT.

В публикации А. Булгаковой [6] рассмотрены проблемы хранения и ис-пользования JWT. Авторы пришли к выводу о том, что хранение токенов в cookies и Local Storage небезопасно, поскольку в этом случае клиент может быть подвергнут атакам типа XSS и CSRF. Был предложен подход с использованием refresh-токенов для обеспечения безопасности пользователей. Но в то же время в статье игнорируются возможные уязвимости на стороне сервера при обработке JWT, а также атаки на сам токен.

Похожая проблема была исследована коллективом ученых [7] на специально разработанном стенде. По итогам эксперимента авторы предлагают ис-пользовать HTTP Only Cookie для предотвращения CSRF-атак на пользователя. Подобная мера позволяет нейтрализовать данную угрозу, но для обеспечения комплексной безопасности системы аутентификации требуется предотвратить атаки не только на клиента, но и на сервер и сам токен.

В работе С.Н. Девициной [8] был проведен анализ проблем, связанных с вопросами аутентификации пользователей на сервисах, продемонстрированы особенности аутентификации и авторизации на основе JWT. Авторами предлагается использование технической информации о клиенте для повышения безопасности взаимодействия с его токеном, что должно предотвратить случаи перехвата JWT злоумышленником. Но для комплексного подхода к данной проблеме требуется убедиться в том, что производится корректная проверка подлинности токена на сервере, и что JWT не содержит иных уязвимостей, позволяющих обойти предложенный авторами подход.

О.Р. Никитин и А.Г. Уймин [9] в своей работе провели тестирование JWT на безопасность, реализовали атаки с использованием алгоритма «none» и удалением сигнатуры. Авторы подчеркивают, что базовой защиты JWT недостаточно для его использования в приложениях и в качестве методов защиты предлагают маркирование трафика на стороне клиента, а также применение VPN для организации защищенного канала передачи токена до сервера.

Вышеупомянутые научные работы рассматривают лишь некоторые аспекты, связанные с безопасностью JWT. Применение

каждой из предложенных мер по отдельности не может обеспечить безопасность системы аутентификации и авторизации. Для комплексного решения данной проблемы и обеспечения безопасности всего жизненного цикла JWT необходимо изучить, каким образом клиент и сервер взаимодействуют с токеном и на каких этапах этого взаимодействия могут возникнуть уязвимости.

Формулировка проблемы

Несмотря на ряд преимуществ, при использовании JWT возникают про-блемы, связанные с безопасностью. Чаще всего в научной и технической литературе описываются уязвимости, связанные с атаками на клиента [6, 7] и на алгоритмы подписи токена [4, 5, 9]. Это ограничивает область выбора защитных мер при обеспечении безопасности JWT из-за малого числа исследуемых уязвимостей и отсутствия полной картины потенциальных атак. В качестве возможных решений данной проблемы мы видим систематизацию существующих уязвимостей и построение классификации векторов атак на JSON Web Tokens, позволяющей установить взаимосвязь между этапами жизненного цикла токена и уязвимостями. Использование классификации позволит, во-первых, на этапе проектирования информационной системы понять слабые места процесса аутентификации и авторизации и ликвидировать их, во-вторых, выявлять уязвимости уже работающих сервисов и своевременно принимать меры по их нейтрализации, в-третьих, с помощью предложенной классификации возможно подобрать инструментарий для комплексного тестирования безопасности JWT на всех этапах его использования.

Жизненный цикл JSON Web Tokens

На рисунке 1 изображена схема взаимодействия пользователя с приложением с использованием JWT. За выпуск токенов отвечает сервис аутентификации (далее – SAuth), в качестве которого может выступать как специально созданный сервис, так и отдельный модуль в рамках существующих сервисов. Процедура проверки подлинности токенов осуществляется без участия SAuth, валидация происходит при их обработке на стороне целевых сервисов.

Для понимания того, каким образом появляются уязвимости при использовании JWT, рассмотрим процессы аутентификации и авторизации на основе токенов. Для этого опишем жизненный цикл JWT, который состоит из нескольких этапов.

Этап 0. Инициация доступа: пользователь должен подтвердить, что он тот, за кого себя выдает, посылая запрос на целевой сервис. Обработка запроса осуществляется посредством SAuth. Данный этап инициирует жизненный цикл JWT.

Этап 1. Получение токена: в случае, если предыдущий этап прошел успешно, SAuth генерирует для пользователя токен. На этом этапе в JWT формируются поля полезной на-

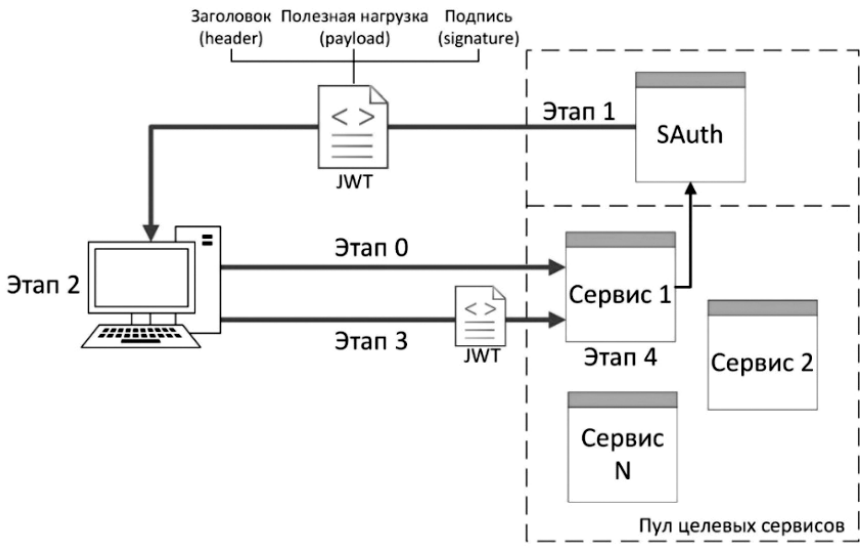

Рис. 1. Описание жизненного цикла JSON Web Token

грузки, в которых, как правило, содержится служебная информация, а также роль и права пользователя в данном приложении. Количество полей токена и их содержание может быть произвольным и определяется администратором системы. Сгенерированный токен может использоваться с целью автоматической сквозной аутентификации и авторизации для всех сервисов, а не только для целевого. После того, как JWT был сформирован, он отправляется клиенту. В случае, если пользователь изначально ввел некорректные данные, возвращается соответствующая ошибка.

Этап 2. Сохранение токена: после получения JWT сохраняется браузе-ром клиента в локальных хранилищах (например, cookies или Local Storage) и используется при дальнейшей аутентификации.

Этап 3. Запрос на аутентификацию: при каждом клиентском обращении к приложению JWT из хранилища автоматически помещается в заголовок http-запроса. При этом должна быть обеспечена безопасная среда передачи токена, поскольку содержащаяся в JWT информация передается в открытом виде.

Этап 4. Обработка токена: при получении JWT сервисом запускается процесс его валидации (за счет проверки подписи токена), параметры которого определяются администратором системы. В случае, если JWT не проходит проверку сервиса, пользователю придется заново пройти этап инициации доступа. Если же проверка пройдена успешно, то сервис, обрабатывая поля токена, производит авторизацию клиента и предоставляет доступ к запрашиваемому ресурсу в соответствии с ролью пользователя в приложении.

Структура JWT

Для понимания природы возможных уязвимостей на всех этапах жизненного цикла токена отдельного рассмотрения требует структура JWT.

JSON Web Token представляет из себя средство авторизации, аутентификации и безопасной передачи информации в формате JSON между двумя сторонами. Структура, описание, варианты использования и формирования JWT определены в стандарте RFC 7519 JSON Web Token (JWT) [10]. Согласно стандарту, токен состоит из 3-х частей: заголовка, полезной нагрузки и подписи. Каждая из частей кодируется алгоритмом Base64 и разделяется друг от друга точкой.

Заголовок (*header*) является первой частью JWT. В нем, как правило, со-держится информация о типе токена (поле «*typ*») и используемом алгоритме подписи (поле «*alg*») [5]. Но поскольку стандарт RFC 7519 носит лишь реко-мендательный характер, то в заголовок можно включать и иную информацию по желанию администраторов системы. Как правило, значение поля «typ» равно «JWT». Возможные для использования в JWT криптографические алгоритмы подписи описаны в стандарте RFC 7518 JSON Web Algorithms (JWA). Стандартом рекомендованы несколько из них: ES256 и RS256, но существует и множество других доступных для использования алгоритмов [11].

Полезная нагрузка (*payload*) является следующей частью JWT, в которой содержится служебная информация, а также необходимые данные для аутентификации и авторизации пользователя. Обычно в данной части токена находятся такие сведения, как уникальный идентификатор пользователя, его имя и роль в приложении (пользователь, администратор и т.п.). Также существуют определенные в стандарте [10] поля, которые предназначены для предоставления служебной информации, например, времени жизни токена, информации о выпускающем токен сервисе и о получателе данного JWT.

Подпись (*signature*) – третья часть JWT, которая обеспечивает безопас-ность процедур аутентификации и авторизации в приложении. Подпись формируется следующим образом: сначала нужно закодировать с помощью алгоритма Base64 поля заголовка и полезной нагрузки, разделив их между собой точкой. Затем полученная строка хэшируется с помощью алгоритма, заданного в заголовке JWT и с использованием секретного ключа. Полученная строка также кодируется с помощью Base64 и помещается после части полезной нагрузки в подпись.

В результате анализа видно, что структура JWT задана не строго, и администратор системы может по своему усмотрению создавать новые поля или удалять старые в любой части токена.

Для определения и классификации векторов атак на JWT необходимо исследовать актуальные и потенциальные уязвимости, опираясь на структуру токена и этапы его жизненного цикла.

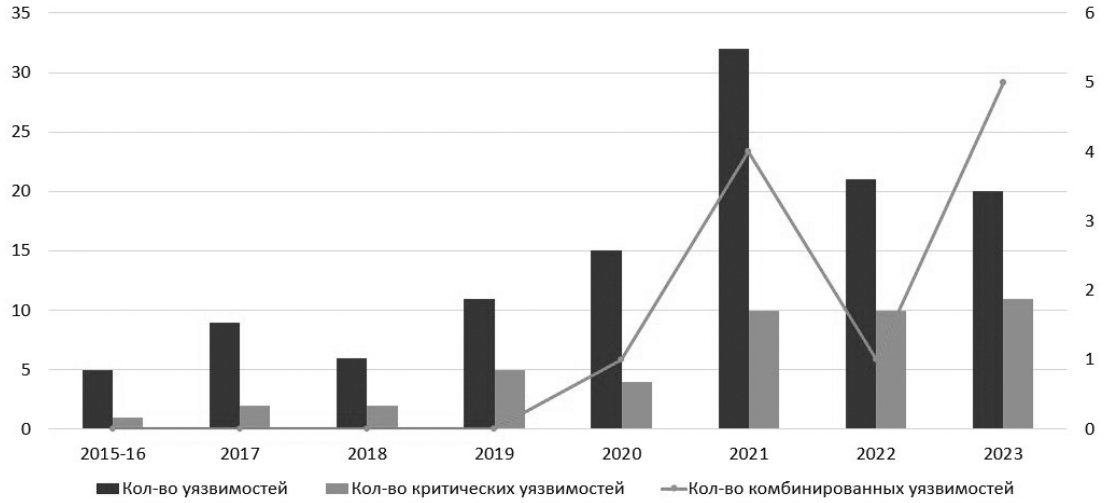

Рис. 2. Диаграмма уязвимостей JWT

Анализ существующих уязвимостей JWT

В ходе исследования нами было проанализировано 119 существующих на момент написания статьи уязвимостей базы данных CVE, которые напрямую связаны с процессами аутентификации на основе JSON Web Tokens. На рисунке 2 представлена динамика появления новых уязвимостей (синий столбец), связанных с JWT, с 2015 по 2023 гг.

Начиная с 2020 г. растет количество критических уязвимостей токенов (серый столбец), т.е. имеющих рейтинг 9.0 и выше по CVSS 3.0, растет и их отношение к общему ежегодному числу новых уязвимостей. При этом с 2021 г. наблюдается снижение количества новых уязвимостей с рейтингом «высокий» (от 7.0. до 8.9) и «средний» (от 4.0. до 6.9) по классификации CVSS 3.0. Оба этих фактора формируют положительную динамику роста среднего значения рейтинга всех новых уязвимостей в год: от 7.7 в 2020 г. до 8.5 в 2023 г.

Все существующие в базе CVE уязвимости JWT можно условно поделить на несколько категорий в зависимости от способов эксплуатации и их локализации:

- уязвимости проверки полей токена;
- уязвимости, связанные с формированием и проверкой подписи JWT;
- уязвимости конфигурации сервера.

Стоит отметить, что с 2021 г. появляются новые категории уязвимостей токенов, которых раньше не было или которые были представлены в единичном экземпляре в 2020 г.:

- уязвимости серверной части, такие как некорректная работа микросервисов и уязвимости в API;
- уязвимости каналов связи, связанные с перехватом токенов или небезопасным процессом передачи JWT;
- уязвимости, которые появляются при использовании OpenSource решений.

Также за этот период в разы увеличилось количество уязвимостей, связанных с формированием и проверкой подписи JWT.

Анализ показал, что 42% уязвимостей, связанных с подписью JWT и с новыми категориями, имеют рейтинг «критический» и составляют 93% от об-щего числа критических уязвимостей за период с 2021 по настоящее время.

Начиная с 2020 г., появляются уязвимости, которые используют слабости одновременно в нескольких областях жизненного цикла JWT и поэтому могут быть отнесены сразу к нескольким категориям. Мы будем называть такие уязвимости комбинированными (оранжевый график). Половина этих уязвимостей являются критическими, а вторая половина имеют рейтинг «высокий».

Подводя итог, мы видим, что с 2021 г. наблюдается резкое увеличение числа уязвимостей JWT. По мнению авторов, это обусловлено массовым использованием токенов в системах аутентификации и авторизации, а также в OpenSource решениях, набравших популярность в постковидный период. При этом за счет использования новых технологий в серверной части (например, микросервисной архитектуры) появляются новые виды

уязвимостей, в том числе комбинированные с уже существующими типами уязвимостей, что существенно повышает итоговый уровень их критичности.

Анализ жизненного цикла JWT

Рассмотрим жизненный цикл JWT с точки зрения возможных уязвимостей. Так как на нулевом этапе создание токена только инициируется, то анализировать жизненный цикл JWT нужно с первого этапа. На данном шаге происходят процессы генерации токена и отправка его клиенту. Выделяются следующие виды возможных уязвимостей:

Хранение в полях токена конфиденциальной информации. В виду того, что JWT хранит информацию в открытом виде, передавать электронные почты, номера телефонов и иную чувствительную информацию пользователей внутри токена недопустимо, поскольку возможен риск компрометации этих данных. Данная уязвимость может эксплуатироваться проведением атак типа «человек посередине» (далее – MITM) при условии использования незащищенного канала связи между клиентом и сервисом, а также проведением XSS-атак (CVE-2022-22311).

Использование криптографически нестойкого секретного ключа. Подпись токена – это гарантия безопасности передаваемой информации и секретный ключ играет здесь главную роль. Если удалось его получить, то будет скомпрометирована вся система аутентификации, а у злоумышленника появится возможность делать запросы к приложению от имени любого пользователя (CVE-2023-26089, CVE-2022-44796, CVE-2022-42980). Данная уязвимость может эксплуатироваться атакой типа bruteforce.

Использование незащищенного канала связи. Если между SAuth и клиентом не обеспечивается защищенный канал связи, становится возможным про-вести атаку MITM. За счет нее нарушитель, находясь в одной сети с пользователем, может перехватывать исходящие и входящие http-пакеты. Если трафик не будет зашифрован, то злоумышленник получает возможность узнать и перехватить токен пользователя в тот момент, когда SAuth будет передавать JWT клиенту (CVE-2022-22311).

На втором этапе жизненного цикла JWT происходит процесс сохранения токена на стороне клиента. На данном шаге существует уязвимость, связанная с *небезопасным хранением токена.*

Сохранение пользовательского JWT возможно в нескольких местах браузера, например, в cookies или LocalStorage. Если токен был сохранен в cookies, то клиент уязвим к атаке типа CSRF [6, 7]. Данная уязвимость влечет за собой возможность выполнения произвольных действий от имени жертвы, а также риск захвата аккаунта (CVE-2020-1762).

Если JWT, полученный от SAuth, сохранятся в LocalStorage браузера, то возникает угроза XSS-атак на пользователя. Их суть состоит в том, чтобы от-править пользователю вредоносную ссылку, при переходе по которой злоумышленник получит информацию из локального хранилища браузера, в котором находится токен жертвы. Итогом эксплуатации XSS-уязвимости является возможность захвата аккаунта пользователя и отправка от его имени произвольных запросов на сервис (CVE-2023-34088, CVE-2021-3509).

На третьем этапе жизненного цикла JWT происходит обращение клиента к серверу с предъявлением токена. Уязвимым в данном случае является канал передачи JWT, и если он не является безопасным, то становится возможным проведение атаки типа MITM.

На последнем этапе цикла выполняется процедура обработки JWT. Дан-ный этап содержит в себе наибольшее количество уязвимостей и возможностей атак на токены:

Некорректная обработка полей JWT. Поскольку стандарт RFC 7519 JSON Web Token (JWT) предусматривает возможность применения произвольных полей в составе JWT, то их нужно правильно и безопасно обрабатывать. Если администраторы системы допустили ошибку при создании механизма обработки полей, то становится возможным проведение целого ряда атак на поля JWT:

1. Поле «username»: обычно оно используются с целью авторизации пользователя на сервисе и определения его роли. Но бывают случаи, когда значение данного поля напрямую подставляется в базу данных с именами пользователей. В таком случае становится возможным провести атаку типа SQL-injection.

2. Поле «role»: как правило, данное поле, если оно присутствует в токене, обозначает, какими привилегиями обладает пользователь. Если значение данного поля не сопоставляется со значениями из других полей токена, то становится возможным повысить привилегии, изменив значения поля, к примеру, на «admin». Также нужно учитывать тот факт, что значения некоторых символов, на-

пример, пробела, могут не учитываться при обработке токена. Таким образом, если значение «admin» вызовет ошибку, то значение «admin» может пройти проверку (CVE-2023-23612).

3. Поле «aud»: является одним из стандартных полей JWT, оно означает, для какого пользователя данный токен был выпущен. Обычно значением данного поля выступает уникальный идентификатор (ID) пользователя. При недостаточной проверке становится возможным повысить привилегии путем смены своего ID на ID администратора (CVE-2018-6873).

4. Поле «iss»: также является стандартным полем и обозначает ресурс, выпустивший данный токен. При проверке данного поля становится возмож-ным изменить его значение так, что JWT будет корректно проходить проверку на тех сервисах, где этого не должно происходить (CVE-2017-8034).

5. Поле «jku» (JWK Set URL): данное поле описано стандартом RFC 7515 JSON Web Signature (JWS) и применятся в случае, когда для формирования подписи токена используются асимметричные алгоритмы шифрования. В данном поле размещается URL-адрес, на котором находится публичный ключ в виде объекта JSON. Если данное поле уязвимо, то становится возможным провести следующую атаку: сначала нужно создать свою пару приватного и публичного ключей, затем добавить на контролируемый веб-сервер файл с расширением .json, в котором будет находиться информация об открытом ключе. После этого необходимо сформировать свой JWT токен и подписать его сгенерированным приватным ключом, а также изменить значение поля «jku» на URL-адрес файла с информацией о публичном ключе. В случае успеха данной атаки, вся система аутентификации и авторизации становится скомпрометированной, поскольку злоумышленник будет способен выдать себя за любого пользователя.

6. Поле «jwk» (JSON Web Key): как и поле «jku», оно описано стандартом RFC 7515 и применяется только при использовании асимметричного алгоритма формирования подписи JWT. Поле предназначено для размещения информации о публичном ключе шифрования. Суть атаки заключается в формировании собственной пары приватный-публичный ключ, размещении информации о публичном ключе в поле «jwk» и подписи токена сформированным приватным ключом. Итог проведения успешной атаки – компрометация системы авторизации и аутентификации (CVE-2019-1010263, CVE-2018-0114).

7. Поле «kid» (Key ID): данное поле является стандартным [12] и предназначено для идентификации ключа, которым производится подпись JWT. Если значение данного поля не подвергать проверке, то оно потенциально уязвимо к таким атакам, как Path Traversal, SQL-injection, Remote Code Execution (RCE).

8. Поле «exp» (Expiration Time) определено стандартом RFC 7519 и предназначено для установления времени, после которого JWT считается недействительным. В случае уязвимости данного поля, токены с истекшим сроком годности считаются действительными и успешно обрабатываются сервисом, что позволяет злоумышленнику действовать от имени другого пользователя неограниченное количество времени (CVE-2020-26892).

Некорректная проверка подписи JWT. Алгоритмы формирования подписи различны и, несмотря на наличие рекомендаций со стороны стандарта RFC 7518, выбор алгоритма зависит только от администратора системы. На этапе проверки подписи возможны следующие виды атак:

1. Использование непредусмотренных алгоритмов: алгоритм формирования подписи задается сервисом аутентификации на этапе генерации токена и находится в поле «alg». Но если на стороне сервисов данное поле не проверяется или проверяется некорректно, становится возможно реализовать несколько атак. Например, в качестве значения поля «alg» можно указать «none» и в случае успешной реализации данной атаки целевой сервис не будет проверять подпись токена, т.е. станет возможным менять содержимое токена произвольным образом (CVE-2022-23540, CVE-2021-22160). Если для формирования подписи используется асимметричный алгоритм шифрования, то можно попробовать провести атаку типа algorithm confusion («путаница в алгоритмах»). Ее суть заключается в том, чтобы подписать токен открытым ключом шифрования и поменять в поле «alg» алгоритм на симметричный. Таким образом, при проверке JWT сервис не выдаст ошибку. Это произойдет потому, что в JWT мы указали симметричный алгоритм шифрования и сервис произвел расшифровку JWT тем ключом, которым мы его зашифровали (CVE-2022-23541, CVE-2021-46743). Итогом реализации

любой из атак является компрометация системы аутентификации и авторизации.

2. Недостаточная проверка сигнатуры токена или ее отсутствие: система проверки токена может быть спроектирована таким образом, что подпись JWT не проверяется или сервис считает действительными токены с нулевой длиной подписи. Это дает возможность любому пользователю выдать себя за другого и действовать от его имени (CVE-2023-2827, CVE-2022-39366, CVE-2022-39227).

Уязвимость механизма отзыва токена: если пользовательский JWT был скомпрометирован, то должен быть механизм прекращения действия данного токена, в противном случае злоумышленник сможет на протяжении длительного времени действовать от имени жертвы (CVE-2022-22332, CVE-2021-35342).

Незащищенность от внутреннего нарушителя: при обеспечении ком-плексной безопасности системы аутентификации и моделировании возможных угроз следует рассматривать как потенциальных злоумышленников не только внешних пользователей, но и тех, кто напрямую имеет доступ к серверному оборудованию и программному обеспечению. Характерными для внутреннего нарушителя атаками можно считать компрометацию секретного ключа для подписи JWT на серверной части (CVE-2021-23207, CVE-2021-3167), а также возможность проведения атаки типа MITM во время отправки токена от SAuth к клиенту.

Хранение секретных ключей в исходном коде приложения: если приложение имеет закрытый исходный код, то риск компрометации секретного ключа для подписи JWT существует только со стороны разработчиков данного сервиса или администраторов системы, если же программное обеспечение является открытым, то любой человек может узнать секретный ключ (CVE-2023-39846, CVE-2023-33372, CVE-2023-33371).

Некорректная обработка входящих запросов: вследствие уязвимости алгоритма обработки пользовательские запросы с некорректной подписью токена вызывают отказ в обслуживании. Зная об этом, злоумышленники могут на некоторое время вывести сервис из строя (CVE-2023-34429, CVE-2021-43824). При неправильной конфигурации сервиса, становится возможным обойти процедуру аутентификации, используя специальные поля в http-заголовках. Данная атака позволяет получить злоумышленнику несанкционированный доступ к сервису (CVE-2023-30845, CVE-2023-27487). Также при получении токена с неправильной подписью, сервис может в http-заголовках ответа прислать значение правильного секретного ключа (CVE-2023-40171, CVE-2022-39304).

Классификация атак на JWT

На основании проведенного анализа этапов жизненного цикла JWT и уязвимостей базы CVE, мы можем выделить классы атак на токены, которые представлены на рисунке 3.

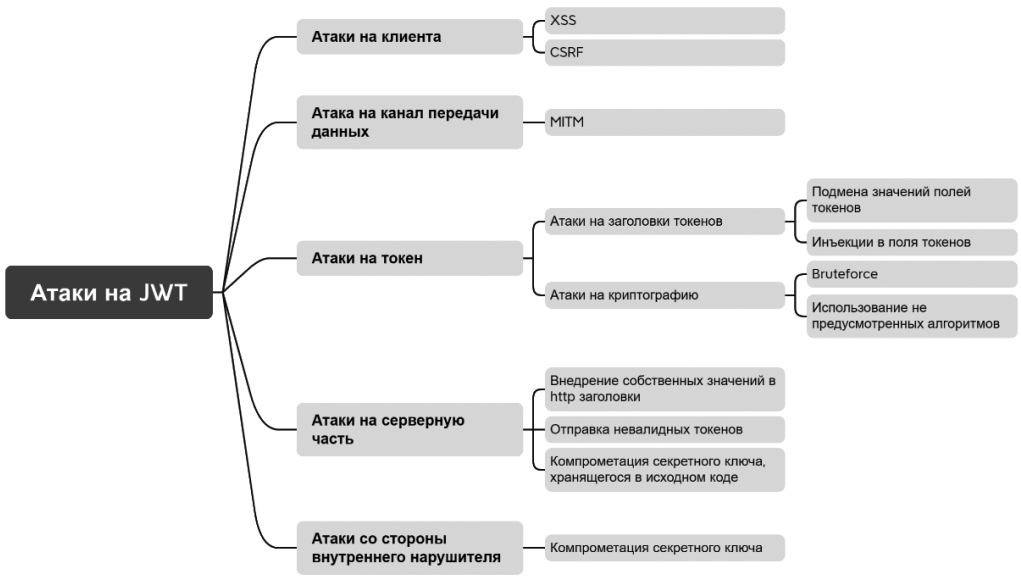

Рис. 3. Классификация атак на JWT

В ходе анализа мы выяснили, что нескольким уязвимостям на различных этапах жизненного цикла JWT может соответствовать одна атака и, как следствие, один объект атаки. В связи с чем было принято решение составить классификацию на основании возможных объектов атак, а также с учетом возможности наличия внутреннего нарушителя.

Атаки на клиента. Данная категория подразумевает под собой возможность эксплуатации уязвимостей посредством проведения атак типа XSS и CSRF, связанных прежде всего с небезопасным хранением JWT.

Атаки на канал передачи данных возможны в результате передачи JWT по незащищенному каналу связи. В связи с тем, что данные передаются в открытом виде, то для получения доступа к содержимому токена будет проведе-ния атаки типа MITM.

Атаки на токен реализуются за счет уязвимостей в структуре JWT. Из-за отсутствия в JWT ограничения на число создаваемых полей могут возникнуть проблемы их корректной и безопасной обработки, что позволяет скомпрометировать систему аутентификации за счет внедрения инъекций в поля заголовков или подмены их значений. Использование некриптосткойкого секретного ключа при подписывании токена и отсутствие контроля алгоритмов позволяют злоумышленникам провести атаки типа bruteforce или algorithm confusion.

Атаки на серверную часть становятся возможны за счет наличия уязви-мостей в конфигурации сервера. Так, например, уязвимость алгоритма обработки токенов с некорректной подписью может привести к ошибке типа отказ в обслуживании, некорректная обработка http-пакетов может позволить обойти механизм аутентификации, а хранение секретного ключа в исходном коде приложения дает возможность действовать от имени любого пользователя данного сервиса.

Атака со стороны внутреннего нарушителя подразумевает вредоносное воздействие на систему со стороны лица, у которого есть непосредственный доступ к серверному оборудованию и/или программному обеспечению. Для внутреннего злоумышленника характерны те же атаки, что и для внешнего, за исключением возможности компрометации секретного ключа при физическом доступе к серверу.

Данная классификация позволяет определить возможные объекты атаки и в сочетании с описанием уязвимостей жизненного цикла JWT позволит упростить построение векторов атак на токены. Как следствие, это обеспечит системный подход к нейтрализации угроз при проектировании информационной системы, использующей аутентификацию и авторизацию на основе JWT, а также поможет при проведении аудитов информационной безопасности. Однозначное разграничение атак и сопоставление их с выявленными уязвимостями позволит более точно подобрать инструментарий для тестирования безопасности JWT.

Заключение

В ходе исследования был проведен анализ существующих 119 уязвимостей базы CVE, связанных с использованием JWT в системе аутентификации. В результате был выявлен резкий рост числа появления новых уязвимостей, особенно критического уровня, начиная с 2021г. Отмечено появление новых категорий уязвимостей, в том числе комбинированных, которые могут быть обусловлены распространением JWT в качестве способа аутентификации и использованием его при работе в микросервисной архитектуре и OpenSource решениях.

Авторами был определен жизненный цикл JSON Web Tokens, состоящий из 4 основных этапов, и описана его структура с учетом анализа стандартов по JWT. Это позволило выявить 11 типов уязвимостей и сгруппировать их в соответствии с принадлежностью к тому или иному этапу жизненного цикла. Некоторые из выявленных уязвимостей не были представлены в научной литературе на данный момент. Для каждого из рассмотренных типов уязвимостей были определены возможные классы атак и способы их реализации. В результате была предложена классификация векторов атак на токены, основанная на их объектах атаки.

Для реализации комплексного подхода к обеспечению безопасности систем аутентификации и авторизации на основе JWT необходимо расширить и дополнить предложенную классификацию с целью обоснованного выбора методов и средств выявления уязвимостей токенов, оценки последствий их эксплуатации, построения векторов атак и определения защитных мер.

Литература

1. Mike Loukides, Steve Swoyer. Microservices Adoption in 2020. URL: https://www.oreilly.com/radar/microservices-adoption-in-2020 (дата обращения: 10.09.2023).

2. Феоктистов И.В. Сравнительное исследование методов аутентифика-ции в информационных системах // Инновации и инвестиции. 2023. № 7. С. 193–198.

3. Kalaska R., Czarnul P. Benchmarking Scalability and Security Configura-tion Impact for A Distributed Sensors-Server IOT Use Case, Proceedings of the 37th International Business Information Management Association (IBIMA), 30–31 May 2021, Cordoba, Spain, ISBN: 978-0-9998551-6-4, ISSN: 2767-9640.

4. Бетелин А.Б., Егорычев И.Б., Прилипко А.А. О некоторых особенно-стях JWT аутентификации в веб-приложениях // Труды научно-исследовательского института системных исследований российской академии наук. 2021. С. 4–10.

5. Mahindrakar P., Pujeri U. Insights of JSON Web Token // International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-6, March 2020. P. 1707–1710. DOI:10.35940/ijrte.F7689.038620.

6. Bulgakova O., Mashkov V., Zosimov V. Risk of Information Loss Using JWT Token // CIT-Risk'2021: 2nd International Workshop on Computational & Information Technologies for Risk-Informed Systems, September 16–17, 2021, Kherson, Ukraine.

7. Darmawan I., Gunawan R., Pramesti D. JSON Web Token Penetration Testing on Cookie Storage with CSRF Techniques // International Conference Advancement in Data Science, E-learning and Information Systems ICADEIS 2021. DOI: 10.1109/ICADEIS52521.2021.9701965.

8. Девицына С.Н., Пилькевич П.В., Удод Е.В. Способы улучшения защищенности сервисов, использующих JWT-токены // Экономика. Информатика. 2023. № 1. С. 144–151.

9. Никитин О.Р., Уймин А.Г. Инфраструктура JSON Web Token. Реализация основных типов атак // ПЕРСПЕКТИВЫ НАУКИ. 2023. № 2. С. 28-34.

10. RFC 7519 JSON Web Token (JWT). URL: https://datatracker.ietf.org/doc/html/rfc7519 (дата обращения: 15.09.2023)

11. RFC 7518 JSON Web Algorithms (JWA). URL: https://datatracker.ietf.org/doc/html/rfc7518 (дата обращения: 15.09.2023)

12. RFC 7515 JSON Web Signature (JWS). URL: https://datatracker.ietf.org/doc/html/rfc7515 (дата обращения: 17.09.2023)

References

1. Mike Loukides, Steve Swoyer. Microservices Adoption in 2020. Available at: https://www.oreilly.com/radar/microservices-adoption-in-2020 (accessed 10 Sep-tember 2023).

2. Feoktistov I.V. Sravnitel'noe issledovanie metodov autentifika-cii v informacionnyh sistemah [Innovacii i investicii]. 2023, no. 7, pp. 193–198.

3. Kalaska R., Czarnul P. Benchmarking Scalability and Security Configura-tion Impact for A Distributed Sensors-Server IOT Use Case, Proceedings of the 37th International Business Information Management Association (IBIMA), 30–31 May 2021, Cordoba, Spain, ISBN: 978-0-9998551-6-4, ISSN: 2767-9640.

4. Betelin A.B., Egorychev I.B., Prilipko A.A. O nekotoryh osobennostjah JWT autentifikacii v veb-prilozhenijah [Trudy nauchno-issledovatel'skogo instituta sistemnyh issledovanij rossijskoj akademii nauk]. 2021, pp. 4–10.

5. Mahindrakar P., Pujeri U. Insights of JSON Web Token [International Journal of Recent Technology and Engineering (IJRTE)]. 2020, vol. 8, issue 6, pp. 1707–1710. DOI:10.35940/ijrte.F7689.038620.

6. Bulgakova O., Mashkov V., Zosimov V. Risk of Information Loss Using JWT Token [CIT-Risk'2021: 2nd International Workshop on Computational & In-formation Technologies for Risk-Informed Systems]. September 16–17 2021, Kherson, Ukraine.

7. Darmawan I., Gunawan R., Pramesti D. JSON Web Token Penetration Testing on Cookie Storage with CSRF Techniques [International Conference Advancement in Data Science, E-learning and Information Systems ICADEIS]. 2021. DOI: 10.1109/ICADEIS52521.2021.9701965.

8. Devicyna S.N., Pil'kevich P.V., Udod E.V. Sposoby uluchshenija zashhishhjonnosti servisov, ispol'zujushhih JWT-tokeny [Jekonomika. Informatika]. 2023, no. 1. pp. 144-151.

9. Nikitin O.R., Ujmin A.G. Infrastruktura JSON Web Token. Realizacija osnovnyh tipov atak [PERSPEKTIVY NAUKI]. 2023, no. 2, pp. 28-34.

10. RFC 7519 JSON Web Token (JWT). URL: https://datatracker.ietf.org/doc/html/rfc7519 (accessed 15 September 2023)

11. RFC 7518 JSON Web Algorithms (JWA). URL: https://datatracker.ietf.org/doc/html/rfc7518 (accessed 15 September 2023)

12. RFC 7515 JSON Web Signature (JWS). URL: https://datatracker.ietf.org/doc/html/rfc7515 (accessed 15 September 2023)

ЗУЛЬКАРНЕЕВ Искандер Рашитович, доцент кафедры информационной безопасности федерального государственного автономного образовательного учреждения высшего образования «Тюменский государственный университет». 625003, г. Тюмень, ул. Володарского, 6. E-mail: i.r.zulkarneev@utmn.ru

БАСАЛАЙ Константин Алексеевич, студент федерального государст-венного автономного образовательного учреждения высшего образования «Тюменский государственный университет». 625003, г. Тюмень, ул. Володар-ского, 6. E-mail: stud0000272406@study.utmn.ru

ZULKARNEEV Iskander Rashitovich, Associate Professor of the Depart-ment of Information Security at the Federal State Autonomous Educational Institu-tion of Higher Education «Tyumen State University». 625003, Tyumen, st. Volodarskogo, 6. E-mail: i.r.zulkarneev@utmn.ru

BASALAY Konstantin Alekseevich, student of the Federal State Autono-mous Educational Institution of Higher Education «Tyumen State University». 625003, Tyumen, st. Volodarskogo, 6. E-mail: stud0000272406@study.utmn.ru

УДК 004.056.5 Вестник УрФО № 2(52) / 2024, с. 89–98

Поршнев С. В., Пономарева О. А. DOI: 10.14529/secur240209

ПРОБЛЕМЫ РАЗРАБОТКИ МОДЕЛИ УГРОЗ ДЛЯ ХРАНИЛИЩА ГЕТЕРОГЕННЫХ ДАННЫХ

В статье на примере данных ограниченного доступа металлургического производства, имеющих разнородную динамически изменяющуюся структуру, для хранения которых используется хранилище гетерогенных данных (ХГД) Автоматизированной Системы Выпуска Металлургической Продукции (АС ВМП) обсуждаются проблемы обеспечения информационной безопасности хранилищ данных данного типа и возможные подход к их решению. Для этого рассмотрена структура АС ВМП; технологии и структура системы управления ХГД МП; проведен анализ используемых в настоящее время мер защиты информации в АС ВМП; определены угрозы информационной безопасности ХГД МП, не учтенные на этапе проектирования АС ВМП.

Обоснована необходимость изменения концепции обеспечения информационной безопасности (ИБ) хранилищ данных, которая предусматривает переход от принятия мер для закрытия известных уязвимостей к созданию защищенных ХГД.

Ключевые слова: *хранилище гетерогенных данных, защищенное хранилище гетерогенных данных, концепция конфиденциальности, целостности и доступности данных, уязвимости хранилищ больших данных, металлургическое производство.*

Porshnev S. V., Ponomareva O. A.

PROBLEMS OF DEVELOPING A THREAT MODEL FOR HETEROGENEOUS DATA STORAGE CONTROL NETWORK

In the article, using the example of limited access data from a metallurgical production facility, which has a heterogeneous dynamically changing structure, for the storage of which the heterogeneous data warehouse (HDS) of the Automated System for the Production of Metallurgical Products (AS VMP) is used, the problems of ensuring information security of data storages of this type and possible approaches to their decision. For this purpose, the structure of the AS VMP is considered; technologies and structure of the control system of the HGD MP; an analysis of the information protection measures currently used in VMP automated systems was carried out; threats to the information security of the CGD MP that were not taken into account at the design stage of the VMP AS were identified.

The need to change the concept of providing information security for data warehouses is substantiated, which provides for a transition from taking measures to close known vulnerabilities to creating secure data storage systems.

Keywords: *heterogeneous data storage, secure storage of heterogeneous data, concept of confidentiality, integrity and availability of data, big data storage vulnerabilities, metallurgical production.*

Введение

Сегодня на каждом промышленном предприятии для управления бизнес-процессами и их оптимизации традиционно используют следующую иерархически выстроенную совокупность информационных систем (ИС): ERP-систему, используемую для решения задач управления предприятием; несколько различных MES-систем, обеспечивающих сбор, анализ и управление производственными процессами на соответствующих этапах жизненного цикла (ЖЦ) выпускаемой продукции, а также различные автоматизированные системы управления технологическими процессами (АСУ ТП), используемые для управления конкретными технологическими процессами и фиксации их параметров. В результате совокупная информация о единице продукции (ЕП) промышленного производства, представляет собой множество разнородных данных, описываемых различными математическими моделями, которые размещаются в различных автономных хранилищах данных (ХД) перечисленных выше ИС.

При расследовании причин выпуска бракованной продукции, закономерно, возникает нетривиальная задача поиска релевантной запросу информации, размешенной одновременно в нескольких автономных ХД. Для ее решения применительно к металлургическому производству (МП) в [7], [8] предложено создать единое ХД, названное хранилищем гетерогенных данных (ХГД), и разместить в нем все описываемые различными информационными моделями данные МП, которые характеризуют состояние единицы продукции (ЕП) МП на каждом из этапов ее жизненного цикла. Для создания ХГД была использована система управления базами данных СУБД ORACLE Stream.

С точки зрения действующих нормативно-правовых актов в области информационной безопасности данные [1]–[6], находящиеся в соответствующих ХД ERP-систем, MES-систем, АСУ, АСУ ТП, в том числе: значения параметров технологических процессов, планы производства, информация о финансовых потоках, о заказчиках и клиентах, представляют собой информацию ограниченного доступа. Значимость и ценность этой информации приводит к необходимости обеспечения защиты элементов информационной инфраструктуры предприятия и самих ХГД.

Отметим, что традиционно каждый производитель СУБД ограничивается развитием концепции конфиденциальности, целостности и доступности данных. В этой связи, предлагаемые производителями СУБД решения в области обеспечения безопасности СУБД предназначены, в основном, на преодоление существующих и уже известных уязвимостей. При этом очевидно, что данный подход обеспечивает решение конкретных задач, но не способствует выработки универсальной концепции безопасности СУБДи ХГД.

В статье на примере АС ВМП проведен анализ современных проблем, особенностей защиты, требований к обеспечению безопасности ХГД МП.

Структура автоматизированной системы выпуска металлургической продукции (АС ВМП)

Автоматизированная Система Выпуска Металлургической Продукции (АС ВМП), обеспечивающая сбор данных, хранение, формирование запросов по слежению, контролю, моделированию, анализу и выдаче рекомендаций по оптимизации как на отдельных этапах производства, так и полного цикла выпуска металлургической продукции была разработана в рамках выполнения договора № 02.G25.31.0055 с Минобрнауки России от 12 февраля 2013 г., срок выполнения с 10.01.2013 по 30.11.2015 г. (проект 2012-218-03-167), заключенного в рамках выполнения Постановления № 218 Правительства Российской Федерации и технические решения изложены в [1, 2].

Из рисунка 1 видно, что АС ВМП состоит из двух взаимосвязанных подсистем: автоматизированной информационной системы сбора и анализа данных (АИС САД) производства и автоматизированной информационной системы моделирования организационной деятельности (АИС МОД) предприятия.

В состав подсистемы АИС САД входят следующие модули: модуль «Хранилище гетерогенных данных металлургического производства» (ХГД МП); модуль «Конструктор запросов» (КЗ);модуль обмена данными с автоматизированными системами предприятия (ОДАСП).

Перечисленные модули выполняют следующие функции: 1) обмен данными между ХГД МП АИС САД и автоматизированными системами МП всех уровней ОДАСП; 2) надежное хранение и быстрый доступ к ХГД МП, поддержка их хронологии, целостности и непротиворечивости, что обеспечивает воз-

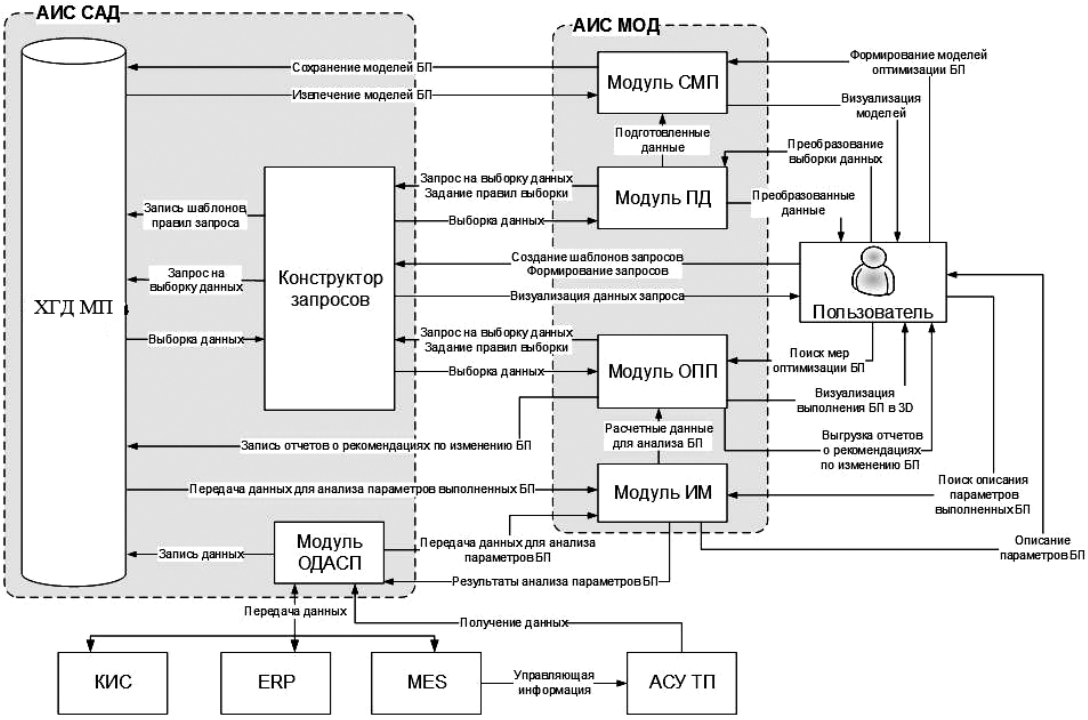

Рис. 1. Структурная схема АС ВМП

можность создания единого информационного пространства данных МП, построение единого интерфейса пользователя, разработка общих алгоритмов обработки данных и осуществление высокопроизводительной аналитической обработки данных; 3) предоставление пользователю удобного интерфейса для создания требуемых выборок информационных параметров и формирование нерегламентированных отчетов (КЗ).

АИС САД обеспечивает сбор, хранение и предоставление информации пользователям, которая может быть использована, в том числе, для создания имитационных моделей технологических, логистических и организационных (бизнес) процессов.

Подсистема АИС МОД состоит из следующих модулей: 1) создания моделей предприятия (СПМ); 2) подготовки данных (ПД); 3) организационных процессов предприятия (ОПП); 4) имитационного моделирования (ИМ).

Модуль ПД предназначен для решения следующих задач: 1) получение выборки данных, сформированной по запросу из модуля КЗ; 2) задание набора правил на обработку выборки данных с помощью графического инструмента создания правил обработки данных; 3) совместное функционирование множества модулей ПД при интеграции в единый программный интерфейс модулей ПД(модулей-источников данных, модулей обработки данных, модулей анализа и модулей верификации); 4) персистентное (долговременное) хранение пользовательских настроек процесса подготовки данных для размещения в ХГД МП; 5) анализ, верификация и преобразование выборки данных в соответствии с заданным набором правил; 6) восстановление пропущенных данных на основе анализа накопленной статистики; 7) анализ данных параметров технологических, логистических и организационных (бизнес) процессов (ТЛОБП) на основе методов машинного обучения (анализ распределённых лагов, фильтры временных рядов, полосовой анализ, методы автоматической классификации данных, нейронные сети) для решения задач прогнозирования и интерпретации подготавливаемых данных.

Модуль СМП реализует следующие функции: 1) создание, модифицирование и запись в ХГД МП моделей ТЛОБП с помощью визуального конструктора моделей мультиагентных имитационных процессов, деревьев анализа параметров процесса, вновь разработанных функциональных элементов для представления графических элементов создания моделей на основе web-интерфейса; 2) оценивание соответствия результатов контрольного

прогона разработанного процесса заданным значениям (отличия не должны превышать 80.%); 3) независимое от пользователя исполнение модели процесса в виде отдельного вычислительного процесса; 4) запуск и одновременное исполнение нескольких экземпляров моделей процессов, создаваемых с помощью модуля СМП; 5) согласование входных и выходных параметров модели процессов, создаваемой с помощью модуля СМП, с параметрами реальных технологических процессов для выполнения модели в модуле ИМ; 6) создание протоколов совещаний, проводимых в рамках реализации типового постоянно действующего бизнес-процесса предприятия по изменению производственных процессов; 7) создание средств обработки обращений пользователей для выполнения обработки инцидентов, связанных с эксплуатацией автоматизированной системы выпуска металлургической продукции (АС ВМП).

Модуль ИМ предназначен для решения следующих задач: 1) получение данных из модуля ОДАСП; получение данных из модуля КЗ; 3) получение данных из модуля ПД; 4) получение описания выполненных на производстве ТЛОБП из модуля ХГД МП; 5) анализ параметров выполненных ТЛОБП (продолжительность анализа – не более 30 мин); 6) выдача результатов анализа в модуль ОДАСП; 7) отображение результатов анализа процессов в зависимости от ролей пользователей; 8) функционирование в режиме разделения процессов выполнения модели и отображения результатов для обеспечения нескольких представлений каждой из выполняющихся моделей; 9) одновременный запуск нескольких модулей ИМ, в каждом из которых выполнение моделей производится независимо друг от друга без средств отображения функционирования модуля ИМ, без ожидания и временных остановок выполняющихся потоков вычислений (в асинхронном неблокирующем режиме выполнения); 10) загрузка имитационной модели в модуль ИМ, созданной в СМП; 11) управление выполняющимися моделями в модуле ИМ с помощью команд на старт, останов и отслеживание состояния.

Модуле ОПП предназначен для решения следующих задач: 1) получение расчётных данных из модуля ИМ; 2) получение выборки данных из модуля КЗ; 3) получение результатов выполнения настроек подготовки данных, созданных в модуле ПД; 4) получение значений параметров выполненных ТЛОБП при взаимодействии с модулем ОДАСП; 5) оптимизация ТЛОБП на основе методов имитационного моделирования (мультиагентного, экспертного, ситуационного моделирования), метода анализа и устранения узких мест ТЛОБП; 6) выдача рекомендаций по изменению технологических, логистических и организационных (бизнес) процессов предприятия; 7) создание отчетов о выполненных процессах предприятия и их запись в ХГД МП, содержащие информацию о времени начала и окончания выполнения процессов предприятия); 8) создание отчетов с рекомендациями по недопущению выявленных инцидентов и вывод отчетов в форматах .doc, .xlsx, .csv, .xml; 9) визуализация выполнения процессов предприятия в формате 3D-анимации; 10) визуальное отображение входных и выходных переменных.

Модуль ХГД МП обеспечивает: 1) централизованное хранение данных АС ВМП; 2) выполнение аналитических запросов КЗ; 3) индексирование данных; 4) кэширование данных; 5) резервное копирование данных АС ВМП.

Для этого в ХГД МП реализованы следующие функции: 1) функция записи параметров ЕП, поступивших от АС ТП, КИС, MES, ERP-систем; 2) функции получения статистических данных; 3) функция создания резервной копии; 4) функция восстановления ХГД МП из резервной копии; 5) вспомогательные функции. В соответствие с ТЗ проекта создания АС ВМП длительность записи данных от: 1)АС ТП должна составлять не более 10 мин; 2)КИС, MES, ERP-систем – не более 30 мин.

Таким образом, модуль ХГД МП осуществляет сбор, хранение и предоставление данных по запросам из других подсистем и/или модулей АС ВМП, в которой предусмотрены следующие сценарии взаимодействия с ХГД МП с другими ее подсистемами и модулями: 1) получение данных из модуля ОДАСП; 2) выборка данных из ХГД МП по запросам модуля КЗ (конструктор запросов) и хранение шаблонов правил запросов; 3) передача данных из ХГД МП в АИС МОД и хранение моделей БП МП.

Для управления ХГД была использована система управления базами данных (СУБД) Oracle Streams [9] и объем хранимых данных составляет около 7 терабайт, выбор которой был определен техническим заданием договора № 02.G25.31.0055.

Технологии и структура системы управления ХГД МП

МП представляет собой сложный многоэтапный процесс: этапа подготовки производства, этап доменного производства, этап сталеплавильного производства, этап прокатного производства. При этом на каждом из этапов происходят как изменение физического состояния единицы продукции (ЕП) МП (шихта – чугун – сталь – прокат), так и, соответственно, информационная модель ЕП МП и ее контент. Описать изменяющуюся структуру информационной модели удается за счет использования кортежей вида «Событие, ЕП, Источник информации». Здесь сущность «событие» обеспечивает создание и хранение контрольных точек производственного процесса, в которых создается фиксируемая информация о ЕП МП (например, начало производства партии продукции, завершение производства партии продукции, отклонение параметром производственного процесса и другие факты), сущность «ЕП» –интегрирует в себе множество разнородных данных, которые собираются в автономных базах данных АСУ ПТ, MES- и ERP-систем, сущность «Источник информации» –множество имен серверов автономных баз данных, которые импортируют данные в АС ВМП.

Схема сущностей, используемых для описания процесса хранения информации о событиях, происходящих в течение жизненного цикла ЕП МП, представлена на рисунке 2 и схема взаимодействия информационных потоков, взаимосвязанных с процессом выпуска ГП МП – на рисунке 3.

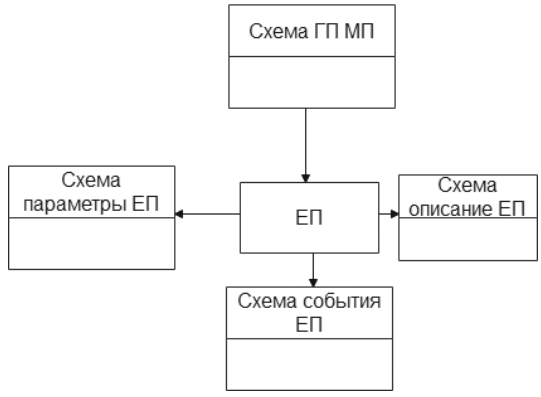

Рис. 2. Схема связей сущностей, используемых для описания схем атрибутов ЕП

Таким образом, основная особенность единого ХГД МП состоит в том, что структура записи данных в единое ХГД МП формируется в момент получения данных от соответствующих источников информации. В этой связи, в едином ХГД МП одновременно с данными необходимо размещать информацию об их структуре, на основе которой далее будет выполняться поиск информации, релевантной запросу пользователя к единому ХГД МП. Применяется технология Oracle Streams, предназначенная для интеграции данных, обмена данными и сообщениями с помощью механизма Advanced Queuing [10] в однородной среде и гетерогенных средах.

Анализ использованных мер защиты информации в АС ВМП

Основными видами воздействий на информационные ресурсы и компоненты информационной системы АС ВМП, которые мо-

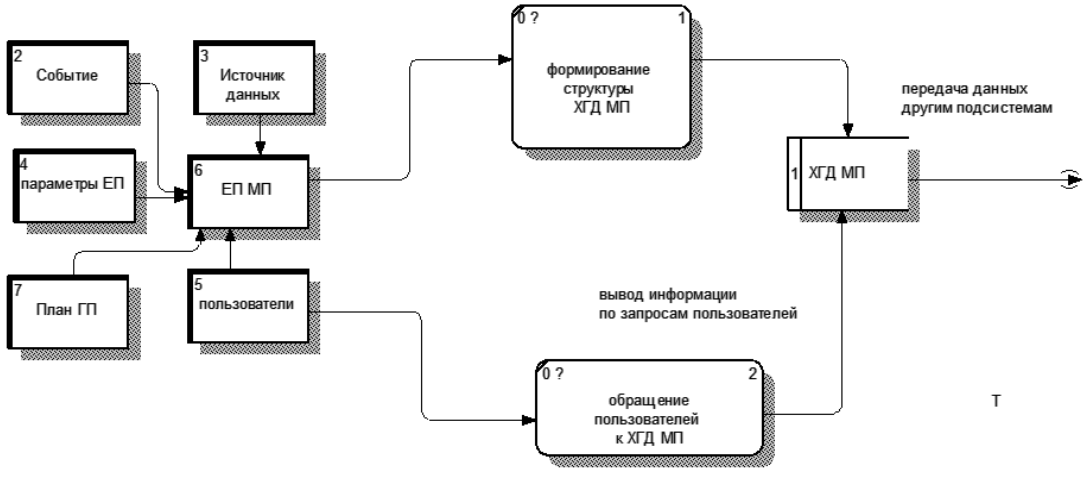

Рисунок 3. Диаграмма потоков данных, размещаемых в едином ХГД МП

гут привести к негативным последствиям, являются: а) утечка (перехват) конфиденциальной информации или отдельных данных (нарушение конфиденциальности); б) несанкционированный доступ к компонентам, защищаемой информации, системным, конфигурационным, иным служебным данным; в) отказ в обслуживании компонентов (нарушение доступности); г) несанкционированная модификация, подмена, искажение защищаемой информации, системных, конфигурационных, иных служебных данных (нарушение целостности); д) несанкционированное использование вычислительных ресурсов в интересах решения несвойственных им задач; е) нарушение функционирования (работоспособности) программно-аппаратных средств обработки, передачи и хранения информации.

В этой связи в техническом задании были предусмотрены следующие меры защиты информации в ХГД МП [1], [11]: 1) выполнение резервного копирования данных; 2) защита от несанкционированного доступа к данным (НСД); 3) обеспечение контроля целостности данных; 4) журналирование (логирования) событий, возникающих в процессе использования ХГД МП.

Выполнение резервного копирования данных

Для реализации функции резервного хранения данных, размещенных в ХГД МП, на языке PL\SQL были реализованы следующие функции: формирование последовательности команд резервного копирования сигнальных данных; резервное копирование оперативных данных; резервное копирование структуры данных; резервное копирование дат, к которым привязаны оперативные данные.

Восстановление данных ХГД МП проходит в следующем строго определенном порядке: актуализация схем данных, восстановление справочных данных, восстановление оперативных данных, восстановление сигнальных данных, актуализация схем.

В процессе актуализации данных ХГД МП: создается начальное состояние схемы базы данных; восстанавливается история изменений структуры данных ХГД МП в виде сохраненных скриптов; восстанавливается схема соответствующей дате (входным параметром для данной процедуры является дата); восстанавливаются только справочные данные, входные параметры dict_data, восстанавливаются сигнальные данные.

Защита от несанкционированного доступа (НСД)

В связи с тем, что ХГД МП реализовано на базе СУБД Oracle, очевидным оказывается решение о применении внутренних механизмов обеспечения защиты от НСД встроенными средствами СУБД, т.к. любая надстройка над СУБД будет работать на более высоком уровне и не обеспечит необходимую степень защиты.

Защита от НСД ХГД МП была реализована с использованием Oracle Database Vault, которая обеспечивает выполнение следующих функций: привилегированный пользовательский доступ к данным приложений; управление доступом к ХГД с помощью многофакторных политик, основанных на заданных условиях; обнаружение и составление отчетов о привилегиях и ролях (матрица доступа), используемых в базе данных; ограничение нерегламентированного доступа к данным приложений за счет предотвращения обхода приложений.

Средства проверки целостности данных

Для проверки целостности данных, находящихся в ХГД МП, использованы соответствующие механизмы проверки и контроля целостности данных, реализованные в СУБД Oracle:средства проверки физической целостности журнала транзакций LGWR (Log Writer) и значения контрольных сумм файлов; средства проверки логической целостности данных (Data Integrity), представленные классическими элементами СУБД Oracle: первичными и внешними ключами (ссылочная целостность), ограничениями (непротиворечивость данных), каскадные удаления.

Журналирование (логирование) событий, возникающих в процессе использования ХГД МП

Журналирование и логирование действий пользователя ХГД МП в системном журнале ХГД МП, а также решение задач мониторинга критический событий ХГД МП (например, проверку объема свободного места на диске, нахождение устаревших индексов), ведение пользовательских логов, осуществляется с помощью функций СУБД Oracle и функций, созданных разработчиком ХГД МП.

Таким образом, при разработке АС ВМП были выполнены требования по обеспечению информационной безопасности в соответствие с техническим заданием (ТЗ), составленного в соответствие с требованиями нормативно-правовой базы в области ИБ на мо-

Таблица 1.

Анализ угроз для ХГД, перечисленных в базе данных угроз ФТЭК России

НомерБДУ	Описание угрозы
УБИ.050: Угроза неверного определения формата входных данных, поступающих в хранилище больших данных	Угроза заключается в возможности искажения информации, сохраняемой в хранилище больших данных, или отказа в проведении сохранения при передаче в него данных в некоторых форматах. Данная угроза обусловлена слабостями технологий определения формата входных данных на основе дополнительной служебной информации (заголовки файлов и сетевых пакетов, расширения файлов и т.п.), а также технологий адаптивного выбора и применения методов обработки мультиформатной (гетерогенной) информации в хранилищах больших данных. Реализация данной угрозы возможна при условии, что дополнительная служебная информация о данных по какой-либо причине не соответствует их фактическому содержимому, или в хранилище больших данных не реализованы методы обработки данных получаемого формата
УБИ.097: Угроза несогласованности правил доступа к большим данным	Угроза заключается в возможности предоставления ошибочного неправомерного доступа к защищаемой информации или, наоборот, возможности отказа в доступе к защищаемой информации легальным пользователям в силу ошибок, допущенных при делегировании им привилегий другими легальными пользователями хранилища больших данных. Данная угроза обусловлена недостаточностью мер по разграничению и согласованию доступа к информации различных пользователей в хранилище больших данных. Реализация данной угрозы возможна при условии использования различных политик безопасности, несогласованных между собой (например, одно средство защиты может отказать в доступе, а другое – предоставить доступ)
УБИ.105: Угроза отказа в загрузке входных данных неизвестного формата хранилищем больших данных	Угроза заключается в возможности отказа хранилищем больших данных в приёме входных данных неизвестного формата от легального пользователя. Данная угроза обусловлена отсутствием в хранилище больших данных механизма самостоятельной (автоматической) адаптации к новым форматам данных. Реализация данной угрозы возможна при условии поступления запроса на загрузку в хранилище входных данных неизвестного формата

мент составления ТЗ (2015 г.) Однако всовременных условиях перечисленных мер по обеспечение ИБ ХГД МПнедостаточно, в связи с возникновением новых угроз [6]. Втаблице 1приведены угрозы информационной безопасности (УБИ), которых не было на момент составления техническогозадания и которые появились в БД ФСТЭК России в настоящее время, актуальные для ХГД МП.

В этой связи, требуется разработка новой концепции обеспечения ИБ ХГД МП, которую можно будет использовать для защиты любых БД, ХД и ХГД.

Обоснование концепции обеспечения ИБ защищенных ХГД

Современные ХД и/или ХГД имеют в своем составе два неотъемлемых друг от друга компонента: данные (собственно ХД и/или ХГД) и программы управления данными. В этой связи, очевидно, что обеспечение безопасности хранимой информации невозможно без обеспечения безопасного управления данными. Исходя из этого, все уязвимости и вопросы безопасности ХД и/или ХГД можно разделить на зависящие от данных и не зависящие от данных. Уязвимости, независящие от данных, являются типичными для любого типа программного обеспечения (ПО). Их причиной, например, может стать несвоевременное обновление ПО, наличие неиспользуемых функций или недостаточная квалификация администраторов ПО. Уязвимости, зависящие от данных, обусловлены, например, тем, что большинство ХД, ХГД поддерживают запросы к данным с помощью того или иного языка запросов, содержащего наборы доступных пользователю функций (которые, в свою очередь, тоже можно считать операторами запросного языка) или произвольные функции на языке программирования. При этом архитектура используемых языков запросов оказывается напрямую связанной с моделью данных, размещаемых в ХД, ХГД.

Следовательно, наличие в ХД, ХГД тех или иных уязвимостей определяется моделью данных. При этом такие уязвимости как, например, инъекции, в зависимости от используемого синтаксиса языка запросов будут реализовываться по-разному (sql-инъекция, java-инъекция).

В связи с выше изложенным, понятно, что, аналогично, можно выделить зависящие и независящие от данных меры обеспечения безопасности хранилищ информации. К независящим от данных можно отнести следующие требования к безопасной системе БД:

1. Функционирование в доверенной среде. (Здесь доверенная средой мы понимаем инфраструктуру предприятия и ее защитные механизмы, предусмотренные к использованию действующими политиками безопасности, то есть в соответствие с правилами безопасности, применяемыми и к каждой из ИС предприятия).

2. Организация физической безопасности файлов данных, которые, в целом, не отличаются от требований, применяемых к любым другим файлам пользователей и приложений.

3. Организация безопасной и актуальной настройки системы управления данными ХД, ХГД, предусматривающая, в том числе, решение общих задач обеспечения безопасности, например, таких как своевременная установка обновлений, отключение неиспользуемых функций или применение эффективной политики паролей.

Следующие требования можно отнести к требованиям, зависящим от данных:

1. Безопасность пользовательского ПО, используемого для управления ХД, ХГД, и доступа к данным, в том числе, построение безопасных интерфейсов и механизмов доступа к данным.

2. Безопасная организация и работа с данными, находящимися в ХД и ХГД, в том числе, обеспечивающая контроль их целостности и доступности (особую актуальность данная задача имеет для ХГД с динамических изменяющейся структурой данных).

Заключение

Для решения обозначенных проблем обеспечения ИБ ХД, ХГД, с нашей точки зрения, целесообразно, перейти от широко используемой сегодня концепции закрытия уязвимостей к концепции комплексного обеспечения их безопасности.

Для практической реализации предложенной концепции необходимо разработать соответствующие методики обеспечения безопасности ХД, ХГД промышленных предприятий, которые позволят избежать ошибок при разработке и реализации организационных и организационно-технических мероприятий в области обеспечения ИБ систем управления ХД и ХГД, а также защититься от наиболее распространенных на сегодняшний день уязвимостей.

Дальнейшие результаты исследований, проводимые авторами в данном направлении (в том числе, разработанные модели угроз и полученные оценки рисков ИБ для ХД, ХГД), являются предметом последующих публикаций.

Литература

1. ГОСТ Р 51275-2006 Объект информатизации. Факторы, воздействующие на информацию. Общие положения.

2. ГОСТ Р 56546-2015 Защита информации. Уязвимости информационных систем. Классификация уязвимостей информационных систем.

3. Требования о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах. Утверждены приказом ФСТЭК России от 15.02.2013 № 17.

4. Состав и содержание организационных и технических мер по обеспечению безопасности информации при их обработке в информационных системах информации с использованием средств криптографической защиты информации, необходимых для выполнения установленных Правительством Российской Федерации требований к защите информации для каждого из уровней защищенности. Утверждены приказом ФСБ России от 10.07.2014 № 378.

5. Методический документ. Методика оценки угроз безопасности информации. Утвержден ФСТЭК России 5 февраля 2021 г.

6. Банк данных угроз безопасности информации ФСТЭК России, URL: https://bdu.fstec.ru/.

7. Porshnev S. The Development of a Heterogeneous MP Data Model Based on the Ontological Approach / S. Porshnev, A. Borodin, O. Ponomareva, S. Mirvoda and O. Chernova // Symmetry. – 2021. – Vol. 13(5). – No. 813.

8. Поршнев С.В., Пономарева О.А. Методология структурного синтеза хранилищ гетерогенных данных промышленного предприятия/ Монография Научно-техническое издательство «Горячая линия – Телеком», Москва, 2022 г.

9. Документация по Oracle Streams URL: https://docs.oracle.com/en/database/oracle/oracle-database/18/strms/index.html.

10. Документация по Advanced Queuing URL: https://docs.oracle.com/en/database/oracle/oracle-database/21/jjdbc/advanced-queuing.html

11. Руководящий документ. Защита от несанкционированного доступа к информации. Термины и определения от 30 марта 1992 года URL: https://normativ.kontur.ru/document?moduleId=1&documentId=198553.

References

12. GOST R 51275-2006 Ob"yekt informatizatsii. Faktory, ozdeystvuyushchiye na informatsiyu. Obshchiye polozheniya. [In Rus]

13. GOST R 56546-2015 Zashchita informatsii. Uyazvimosti informatsionnykh sistem. Klassifikatsiya uyazvimostey informatsionnykh sistem. [In Rus]

14. Trebovaniya o zashchite informatsii, ne sostavlyayushchey gosudarstvennuyu taynu, soderzhashcheysya v gosudarstvennykh informatsionnykh sistemakh. Utverzhdeny prikazom FSTEK Rossii ot 15.02.2013 № 17. [In Rus]

15. Sostav i soderzhaniye organizatsionnykh i tekhnicheskikh mer po obespecheniyu bezopasnosti informatsii pri ikh obrabotke v informatsionnykh sistemakh informatsii s ispol'zovaniyem sredstv kriptograficheskoy zashchity informatsii, neobkhodimykh dlya vypolneniya ustanovlennykh Pravitel'stvom Rossiyskoy Federatsii trebovaniy k zashchite informatsii dlya kazhdogo iz urovney zashchishchennosti. Utverzhdeny prikazom FSB Rossii ot 10.07.2014 № 378. [In Rus]

16. Metodicheskiy dokument. Metodika otsenki ugroz bezopasnosti informatsii. Utverzhden FSTEK Rossii 5 fevralya 2021 g. [In Rus]

17. Bank dannykh ugroz bezopasnosti informatsii FSTEK Rossii, URL: https://bdu.fstec.ru/. [In Rus]

18. Porshnev S. The Development of a Heterogeneous MP Data Model Based on the Ontological Approach / S. Porshnev, A. Borodin, O. Ponomareva, S. Mirvoda and O. Chernova // Symmetry. – 2021. – Vol. 13(5). – No. 813.

19. Porshnev S.V., Ponomareva O.A. "Metodologiya strukturnogo sinteza khranilishch geterogennykh dannykh promyshlennogo predpriyatiya". Monografiya. Nauchno-tekhnicheskoye izdatel'stvo «Goryachaya liniya – Telekom», Moskva, 2022 g. [In Rus]

20. Dokumentatsiya po Oracle Streams URL: https://docs.oracle.com/en/database/oracle/oracle-database/18/strms/index.html. [In Rus]

21. Dokumentatsiya po Advanced Queuing URL: https://docs.oracle.com/en/database/oracle/oracle-database/21/jjdbc/advanced-queuing.html [In Rus]

22. Rukovodyashchiy dokument. Zashchita ot nesanktsionirovannogo dostupa k informatsii. Terminy i opredeleniya ot 30 marta 1992 goda URL: https://normativ.kontur.ru/document?moduleId=1&documentId=198553 [In Rus]

ПОРШНЕВ Сергей Владимирович, доктор технических наук, профессор, директор Учебно-научного центра «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: s.v.porshnev@urfu.ru

ПОНОМАРЕВА Ольга Алексеевна, кандидат технических наук, доцент Учебно-научного центра «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: o.a.ponomareva@urfu.ru

PORSHNEV Sergey Vladimirovich, Doctor of Technical Sciences, Professor, Director of the Educational and Scientific Center «Information Security» of the Federal State Autonomous Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N. Yeltsin». 620002, Yekaterinburg, st. Mira, 32. E-mail: s.v.porshnev@urfu.ru

PONOMAREVA Olga Alekseevna, Candidate of Technical Sciences, assistant professor of the Federal State Autonomous Educational Institution of Higher Education “Ural Federal University named after the first President of Russia B.N. Yeltsin". 620002, Yekaterinburg, st. Mira, 32. E-mail: o.a.ponomareva@urfu.ru

УДК 004.94 Вестник УрФО № 2(52) / 2024, с. 99–105

Воробьев А. П., Кротова Е. Л., Воробьева Е. Ю. DOI: 10.14529/secur240210

ПРЕИМУЩЕСТВА КРИПТОГРАФИИ НА ЭЛЛИПТИЧЕСКИХ КРИВЫХ ДЛЯ РЕШЕНИЯ ЗАДАЧ АУТЕНТИФИКАЦИИ

Цифровые технологии проникают во все сферы нашей жизни, поэтому крайне важным становится вопрос защиты информации и передачи данных. Для обеспечения безопасности, необходимо использовать современные методы криптографии. В данной статье рассматриваются эллиптические кривые, лежащие в основе таких методов, описаны их свойства и особенности выбора. Все более актуальной становится задача противостояния постоянно совершенствующимся злоумышленникам при растущих производительных мощностях при этом вычислительные ресурсы легального пользователя ограничены.

Ключевые слова: *криптография, эллиптические кривые, защита информации.*

Vorobev A. P., Krotova E. L., Vorobeva E. Yu.

ADVANTAGES OF ELLIPTIC CURVE CRYPTOGRAPHY FOR SOLVING AUTHENTICATION PROBLEMS

Digital technologies penetrate into all areas of our lives, so the issue of information protection and data transfer becomes extremely important. To ensure security, it is necessary to use modern cryptography methods. This article discusses the elliptic curves that underlie such methods, describing their properties and selection features. The task of confronting constantly improving attackers with growing production capacities is becoming increasingly urgent, while the computing resources of the legal user are limited.

Keywords: *cryptography, elliptic curves, information security.*

В современных технологиях информационной безопасности широко применяется математический аппарат, основанный на эллиптических кривых. Это связано с их универсальностью, эффективностью и высокой безопасностью [1]. Приведем лишь некоторые примеры областей, в которых он применяется:

Информационная безопасность: для защиты данных и информации в сетях передачи данных, в различных протоколах безопасности, в технологиях и системах аутентификации и авторизации.

Мобильные устройства: для обеспечения безопасности взаимодействий пользователей с приложениями, для шифрования данных и защиты личной информации.

Криптография: для построения современных криптографических методов, таких как эллиптический криптографический обмен ключами, цифровая подпись, блочные шифры.

Финансовые технологии: в различных алгоритмах шифрования платежей, электронных транзакций и других финансовых операций.

Блокчейн технологии: в криптовалютах, например, таких как Биткоин, для обеспечения безопасности транзакций, создания цифровых подписей и других криптографических операций.

Сетевая безопасность: для защиты сетей, обеспечения конфиденциальности данных и предотвращения несанкционированного доступа.

Интернет вещей (IoT): для эффективности и безопасности в устройствах с ограниченными вычислительными ресурсами.

Использование эллиптических кривых позволяет строить криптографические системы с более короткими длинами ключей, но получать при этом эквивалентный уровень безопасности по сравнению с другими методами [2]. Таким образом, криптография на эллиптических кривых может стать решением для устройств с ограниченными вычислительными ресурсами, например, такими как IoT [3].

Рассмотрим конечное поле F_p, характеристика которого $p > 3$.

Множество точек (x, y), координаты которых удовлетворяют уравнению:

$$y^2 + a_1 \cdot xy + a_3 \cdot y = x^3 + a_2 \cdot x^2 + a_4 \cdot x + a_5 \quad (1)$$

называется *эллиптической* кривой. Здесь a_1, a_2, a_3, a_4, a_5, – коэффициенты, принадлежащие полю F_p. В случае характеристики поля, отличной от 2 и 3, уравнение (1) с помощью, подходящей замены переменных может быть приведено к каноническому виду в форме Вейерштрасса:

$$E : y^2 = x^3 + ax + b \quad (2)$$

где $a, b \in F_p$. В качестве поля F_p выбирают чаще всего комплексные ($\mathbb{C}$) или действительные ($\mathbb{R}$) числа.

Наглядно можно представить эллиптические кривые E как линии пересечения поверхности $y^2 = x^3 + ax + b + z$ с плоскостями $z = const$ [4] (см. рис 1).

Эллиптические кривые, получаемые в таких сечениях, в отличие от конических, невозможно параметризировать рациональными функциями.

Свойства эллиптических кривых

1. Так как график кривой E симметричен относительно оси абсцисс, то найти точки пересечения графика с осью Ox можно, решив уравнение:

$$x^3 + ax + b = 0 \quad (3)$$

Для решения таких уравнений третьей степени можно использовать формулы Кардано с вычислением дискриминанта D, равного $E : \left(\frac{a}{3}\right)^3 + \left(\frac{b}{2}\right)^2$. Известно, что при $D<0$ уравнение (3) имеет три различных действительных корня, при $D=0$ два различных (один кратный, кратности 2), а при $D>0$ один действительный и два комплексных сопряженных корня.

Эллиптическая кривая E, заданная уравнением (2), в случае $D=0$ называется сингу-

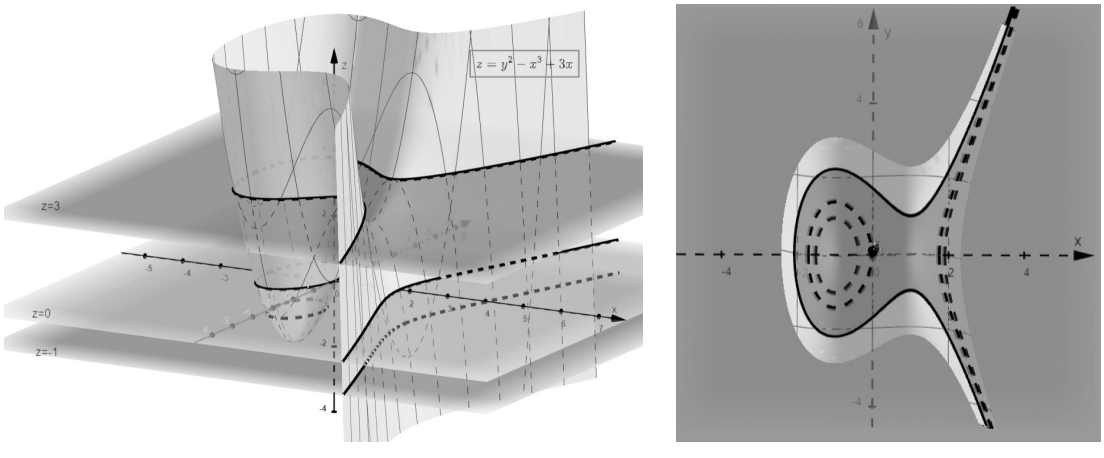

Рис. 1. Пересечение поверхности $z = y^2 - x^3 + 3x$ и плоскостей $z = -1,0,3$

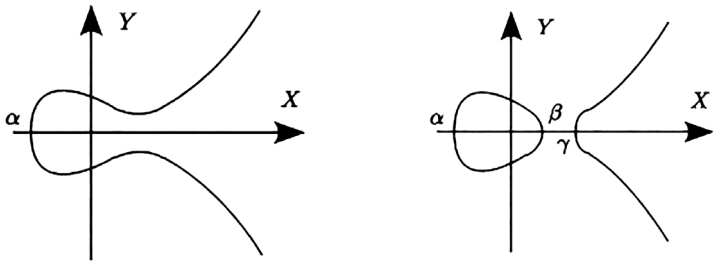

Рис. 2. Примеры несингулярных эллиптических кривых

лярной, а в случае $D \neq 0$ – несингулярной [2]. Нас будут интересовать несингулярные кривые, так как при сингулярности резко снижается криптостойкость алгоритмов. Таким образом, многочлен, стоящий в правой части уравнения (2), не имеет кратных корней и для его коэффициентов выполняется условие

$$\Delta(E) = 4a^3 + 27b^2 \neq 0 \qquad (4)$$

$\Delta(E)$ называется дискриминантом кривой E.

Примеры несингулярных кривых изображены на рисунке 2.

Решения (x, y) уравнения (2) называются афинными точками кривой E. Также в рассмотрение вводится еще одна точка, называемая точкой на бесконечности, и обозначаемая O (нейтральный, нулевой элемент) [5].

2. Введем операцию *сложения* точек на эллиптической кривой.

Отметим, что любая вертикальная прямая l, которая параллельна оси Oy, либо не пересекает кривую E, либо пересекает её дважды. Также, любая другая прямая l пересекает кривую E в одной или в трех точках. Точка касания считается двойным пересечением.

Итак, пусть E – эллиптическая кривая, описываемая уравнением (2), и пусть некая прямая l пересекает E в точках $P(x_1, y_1)$ и $Q(x_2, y_2)$. Тогда эта прямая пересечет кривую E в третьей точке, которую обозначим через $R'(x_3, -y_3)$, (см. рисунок 3). Суммой афинных точек P и Q назовем точку $R(x_3, y_3) = P + Q$, симметричную R' относительно оси Ox. Назовем ее обратной к точке R' и будем обозначать: $R = -R'$.

Если $P = Q$, то прямая l является касательной к кривой E (см. рисунок 3). Будем использовать обозначение: $R = P + P = [2]P$ – удвоение точки P.

В *особых* случаях, распространяя операцию сложения на бесконечно удаленную точку, определим ее так:

- Если l – вертикальная прямая, проходящая через точки P и $Q = -P$, то будем считать, что она пересекает кривую E в бесконечно удаленной точке O, т. е. $P + (-P) = O$.
- Если $P = O$ или $Q = O$, то получаем, что $P + O = P$ или $O + Q = Q$.

Итак,

1) Пусть $P \neq Q$, $P \neq -Q$, $P \neq O$, $Q \neq O$, $P + Q = -R' = R$. Определим координаты точки $R(x_3, y_3)$ [6,7].

Угловой коэффициент прямой l, проходящей через точки $P(x_1, y_1)$ и $Q(x_2, y_2)$, $k = \dfrac{y_2 - y_1}{x_2 - x_1}$. Тогда уравнение прямой l имеет вид: $y = kx + \beta$, где $\beta = y_1 - kx_1$.

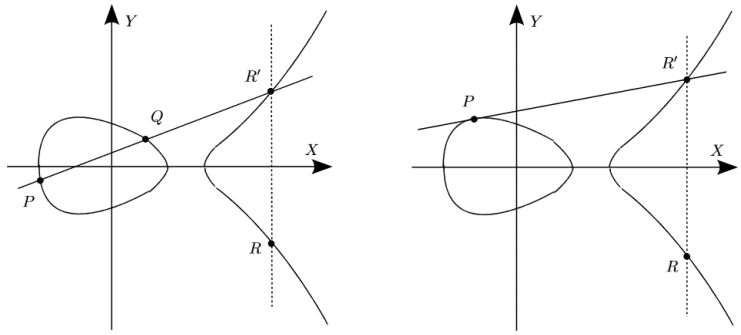

Рис. 3. Сложение точек $P+Q = R$

Точки пересечения прямой l с E являются решениями уравнения:

$$(kx+\beta)^2 - x^3 - ax - b = 0$$

или, учитывая, что оно имеет три различных действительных корня,

$$(kx+\beta)^2 - x^3 - ax - b = -(x-x_1)(x-x_2)(x-x_3) = 0$$

Приравнивая коэффициенты при одинаковых степенях x, получаем:

$k^2 = x_1 + x_2 + x_3$, откуда $x_3 = k^2 - x_1 - x_2$.

С другой стороны, угловой коэффициент k может быть найден как $k = \frac{-y_3 - y_1}{x_3 - x_1}$, тогда имеем: $y_3 = k(x_1 - x_3) - y_1$.

Таким образом, координаты $R(x_3, y_3)$ определяются по формулам:

$$\begin{cases} x_3 = k^2 - x_1 - x_2 \\ y_3 = k(x_1 - x_3) - y_1 \end{cases} \quad (5)$$

2) Пусть $P = Q \neq O$, то есть $R = [2]P$.
Уравнение касательной к кривой $F(x, y) = 0$ в точке $P(x_1, y_1)$ имеет вид:

$$F'_x(P) \cdot (x - x_1) + F'_y(P) \cdot (y - y_1) = 0$$

Тогда угловой коэффициент k касательной определяется по формуле:

$$k = -\frac{F'_x(P)}{F'_y(P)}, \text{ если } F'_y(P) \neq 0.$$

Тогда получаем:

$$k = \frac{3x_1^2 + a}{2y_1} \quad (6)$$

Для определения координат точки $R(x_3, y_3)$ воспользуемся формулами (5) с угловым коэффициентом из (6):

$$\begin{cases} x_3 = k^2 - 2x_1 \\ y_3 = k(x_1 - x_3) - y_1 \end{cases} \quad (7)$$

Если же $F'_y(P) = 0$, то касательная является вертикальной прямой, $[2]P = O$.

Теорема: множество афинных точек кривой E, с нулевым элементом O, является аддитивной абелевой группой относительно операции сложения [5], при этом:

1) $P + Q = Q + P, \ \forall P, Q \in E$,

2) $P + (Q + S) = (P + Q) + S, \ \forall P, Q, S \in E$

3) $\exists$ нулевой элемент O, такой что $P + (-P) = O$,

4) $\forall P \in E \quad \exists$ обратный элемент $-P \in E$: $P + (-P) = O$.

Задача дискретного логарифмирования и выбор параметров эллиптической кривой

В указанных выше формулах (2-7) используются только операции сложения, вычитания, умножение и деления, поэтому при вычислениях с целыми числами по модулю простого числа p все тождества в уравнениях сохранятся, то есть мы находимся в поле вычетов по модулю p.

В результате уравнение (2) преобразуется в:

$$E: y^2 \equiv x^3 + ax + b \pmod{p} \qquad (8)$$

а ограничение (4) примет вид:

$$4a^3 + 27b^2 \neq 0 \pmod{p} \qquad (9)$$

Число $J(E) = 1728 \cdot \dfrac{4a^3}{4a^3 + 27b^2} \pmod{p}$ называется инвариантом эллиптической кривой.

Пусть кривая E определена над простым полем F_p и точка $P \in E(F_p), P \neq O$.

n-кратным точки P назовем композицию (сумму):

$$[n]P = \underbrace{P + P + \ldots + P}_{n \text{ раз}}$$

Наименьшее натуральное число n такое, что $[n]P = O$, называется порядком P в группе $E(F_p)$. Количество точек кривой E будем обозначать N.

Для описания эллиптической кривой в национальном стандарте РФ (ГОСТ Р 34.10-2012) используются следующие параметры [8]:

p – простое число, модуль эллиптической кривой, $p>3$;

a, b – коэффициенты уравнения (8);

P – порождающая точка, или генератор точек кривой, точка большого порядка q. При $1 \le m \le q-1$ последовательность $\{[m]P\}$ задает все различные точки кривой, $[q]P = O$;

q – порядок точки P или порядок подгрупп; $2^{254} < q < 2^{256}$ или $2^{508} < q < 2^{512}$;

$h = \dfrac{N}{q}$ – кофактор кривой [9].

В криптографических системах используются кривые, количество точек которых является большим простым числом. В этом случае любая точка P является генератором всего множества точек. Порождающая точка используется для создания общего открытого ключа и для выполнения операций шифрования и расшифрования. Для безопасности криптосистемы при выборе точки необходимо, чтобы ее порядок был большим простым числом.

Для алгоритмов на эллиптических кривых используются подгруппы с высоким порядком, а значение кофактора h выбирается как можно меньше.

Коэффициенты a и b эллиптической кривой E по известному инварианту $J(E)$ определяются следующим образом:

$$\begin{cases} a \equiv 3k \pmod{p} \\ b \equiv 2k \pmod{p} \end{cases}$$

где $k \equiv \dfrac{J(E)}{1728 - J(E)} \pmod{p}$, $J(E) \neq 0$ или $J(E) \neq 0,1728$.

Задача состоит в нахождении такого числа m, что $[m]P = Q$ при известных P и Q. Такая задача называется задачей *дискретного логарифмирования* и является вычислительно крайне сложной, если тщательно выбирать параметры кривой; во всяком случае, на сегодняшний день не существует известного алгоритма решения такой задачи с полиномиальным временем [3].

Правильный выбор параметров эллиптической кривой для применения современных методов криптографии заслуживает особого внимания. Это поможет обеспечить высокий уровень безопасности данных и защитить их от возможных атак и утечек информации.

Основным достоинством криптосистем на основе эллиптических кривых является то, что они обеспечивают надежность, адекватную классическим криптосистемам RSA или Эль-Гамаля, на значительно меньших по длине ключах, а это существенно сокращает время кодирования и декодирования. Криптосистемы цифровой подписи на основе эллиптических кривых с длиной ключа 160 бит имеют одинаковую стойкость с криптосистемами DSA и Эль-Гемаля с длиной ключа 1024 бита. Электронная цифровая подпись – это эффективное средство защиты информации от модификации, которое переносит свойства реальной подписи под документом в область электронного документооборота. В основу ЭЦП положены такие криптографические методы, как асимметричное шифрование и хэш-функции.

Литература

1. Научные и методологические проблемы информационной безопасности (сборник статей). Под ред. В. П. Шерстюка. — М.: МЦНМО, 2004. — 208 с.

2. Рябко, Б. Я. Криптографические методы защиты информации: учеб. пособие / Б. Я. Рябко, А. Н. Фионов. — 2-е изд., стер. — М.: Горячая линия – Телеком, 2012. — 229 с.

3. Калхиташвили Д. Ш. Операционные системы интернета вещей: возможности, проблемы и решения / Д. Ш. Калхиташвили — Москва: РАНХиГС, 2023. — 5 с.

4. How Elliptic Curve Cryptography Works. [Электронный ресурс]: https://allaboutcircuits.com/technical-articleles/ elliptic-curve-cryptography-in-embedded-systems/

5. Жданов, О. Н. Применение эллиптических кривых в криптографии: учеб. пособие / О. Н. Жданов, Т. А. Чалкин. — Красноярск: СибГАУ, 2011 - 65 с.

6. Эллиптическая криптография: теория/ [Электронный ресурс]: https://habr.com/ru/articles/188958/

7. Elliptic Curve Cryptography: finite fields and discrete logarithms [Электронный ресурс]: https://andrea.corbellini.name/2015/05/23/elliptic-curve-cryptography-finite-fields-and-discrete-logarithms/

8. Национальный стандарт Российской Федерации. Криптографическая защита информации. [Электронный ресурс]: https://www.altell.ru/legislation/standards/gost-34.10-2012.pdf

9. Chandrasekhara, K.R. Elliptic Curve based authenticated session Key establishment protocol for High Security Applications in Constrained Network environment / K.R. Chandrasekhara, M.P. Pillai1 and Sebastian, 2010. [Электронный ресурс]: http://www.arxiv.org/pdf/1202.1895

References

1. Nauchnye i metodologicheskie problemy informacionnoj bezopasnosti (sbornik statej). Pod red. V. P. Sherstjuka. — M.: MCNMO, 2004. — 208 s.

2. Rjabko, B. Ja. Kriptograficheskie metody zashhity informacii: ucheb. posobie / B. Ja. Rjabko, A. N. Fionov. — 2-e izd., ster. — M.: Gorjachaja linija – Telekom, 2012. — 229 s.

3. Kalhitashvili D. Sh. Operacionnye sistemy interneta veshhej: vozmozhnosti, problemy i reshenija / D. Sh. Kalhitashvili — Moskva: RANHiGS, 2023. — 5 s.

4. How Elliptic Curve Cryptography Works. [Электронный ресурс]: https://allaboutcircuits.com/technical-articleles/ elliptic-curve-cryptography-in-embedded-systems/

5. Zhdanov, O. N. Primenenie jellipticheskih krivyh v kriptografii: ucheb. posobie / O. N. Zhdanov, T. A. Chalkin. — Krasnojarsk: SibGAU, 2011 - 65 s.

6. Jellipticheskaja kriptografija: teorija/ [Jelektronnyj resurs]: https://habr.com/ru/articles/188958/

7. Elliptic Curve Cryptography: finite fields and discrete logarithms [Электронный ресурс]: https://andrea.corbellini.name/2015/05/23/elliptic-curve-cryptography-finite-fields-and-discrete-logarithms/

8. Nacional'nyj standart Rossijskoj Federacii. Kriptograficheskaja zashhita informacii. [Jelektronnyj resurs]: https://www.altell.ru/legislation/standards/gost-34.10-2012.pdf

9. Chandrasekhara, K.R. "Elliptic Curve based authenticated session Key establishment protocol for High Security Applications in Constrained Network environment" / K.R. Chandrasekhara, M.P. Pillai1 and Sebastian, 2010. [Jelektronnyj resurs]: http://www.arxiv.org/pdf/1202.1895

ВОРОБЬЕВ Артем Павлович, студент, кафедра Технология твердых химических веществ Казанского национального исследовательского технологического университета. 420015, Республика Татарстан, Казань, ул. Карла Маркса, 68. E-mail: drsleepwalker@yandex.ru

КРОТОВА Елена Львовна, кандидат физико-математических наук, доцент, кафедра Высшей математики, Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: lenkakrotova@yandex.ru

ВОРОБЬЕВА Елена Юрьевна, старший преподаватель, кафедра Прикладной математики, Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: lena-vorobey@yandex.ru

VOROBEV Artem Pavlovich, student, Department of Solid Chemical Substances Technology, Kazan National Research Technological University. 420015, Republic of Tatarstan, Kazan, Karl Marx St., 68. Email: drsleepwalker@yandex.ru

KROTOVA Elena Lvovna, Candidate of Physical and Mathematical Sciences, Associate Professor, Department of Higher Mathematics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. Email: lenkakrotova@yandex.ru

VOROBEVA Elena Yuryevna, Senior Lecturer, Department of Applied Mathematics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. Email: lena-vorobey@yandex.ru

УДК: 004.93'1 Вестник УрФО № 2(52) / 2024, с. 106–121

Панфилова И. Е., Ложников П. С. DOI: 10.14529/secur240211

ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ НЕЙРОСЕТЕВЫХ ПРЕОБРАЗОВАТЕЛЕЙ «БИОМЕТРИЯ-КОД» ДЛЯ ЗАДАЧИ ОБНАРУЖЕНИЯ АТАК НА БИОМЕТРИЧЕСКОЕ ПРЕДЪЯВЛЕНИЕ[1]

В статье рассматривается применение нейросетевых преобразователей «биометрия-код» (НПБК) для защиты блоков обнаружения спуфинг-атак (атак на биометрическое предъявление) в системах биометрической аутентификации по лицу. Введение эффективных антиспуфинг методов имеет критическое значение для повышения надежности биометрических систем, однако внедрение таких методов может создать новые уязвимости. Предложенное решение основано на интеграции НПБК с глубокой нейронной сетью для осуществления бинарной классификации входных изображений лиц на реальные и поддельные. Это обеспечивает безопасное связывание биометрических данных с криптографическими ключами, снижая вероятность несанкционированного доступа. Экспериментальные результаты, полученные на наборе данных CelebA-Spoof, демонстрируют высокую точность работы модуля (97,2%) и низкий уровень средней ошибки классификации (ACER = 2,9%), что подтверждает эффективность предложенного подхода в обеспечении высокой защищенности процедуры аутентификации.

Ключевые слова: *нейросетевые преобразователи, глубокие нейронные сети, атака на биометрическое предъявление, спуфинг атака, распознавание лиц, биометрическая аутентификация.*

[1] Исследование выполнено при финансовой поддержке Минцифры России (Грант ИБ), проект № 40469-18/23-К

AN INVESTIGATION THE APPLICATION OF FUZZY NEURAL EXTRACTORS FOR FACE ANTI-SPOOFING

The article discusses the application of fuzzy neural extractors to protect spoofing attack detection units in facial biometric authentication systems. The introduction of effective anti-spoofing techniques is critical to improve the reliability of biometric systems, but the implementation of such techniques may create new vulnerabilities. The proposed solution is based on the integration of fuzzy neural extractors with a deep neural network to perform binary classification of input face images into real and spoofed ones. This provides secure binding of biometric data with cryptographic keys, reducing the probability of unauthorized access. Experimental results obtained on the CelebA-Spoof dataset demonstrate high accuracy of the module (97.2%) and low level of average classification error (ACER = 2.9%), which confirms the effectiveness of the proposed approach in ensuring high security of the authentication procedure.

Keywords: *neural network converters, deep neural networks, biometric presentation attack, spoofing attack, face recognition, biometric authentication..*

Введение

Системы распознавания лиц представляют собой технологические решения, способные идентифицировать и/или верифицировать личность человека с помощью его лицевых характеристик. Эти системы нашли широкое применение в различных областях, таких как финансы, безопасность, здравоохранение и информационные технологии. Несмотря на высокую степень надёжности систем биометрической аутентификации по лицу, они подвержены угрозам, связанным с подделкой биометрических данных – спуфинг атакам или атакам на биометрическое предъявление (согласно ГОСТ Р 58624.1-2019 [1]). Разработка и внедрение эффективных методов противодействия спуфинг атакам имеет критическое значение для обеспечения доверия к биометрическим системам аутентификации и защиты конфиденциальной информации (персональных данных пользователей).

Современные исследования в области антиспуфинга направлены на создание алгоритмов, способных надёжно классифицировать реальные и поддельные лица, используя передовые технологии глубокого обучения. Однако добавление решений антиспуфинга в системы биометрической аутентификации по лицу способны спровоцировать появление новых уязвимостей системы распознавания: на этапе обнаружения поддельных изображений может быть осуществлена кража биометрического шаблона. В этой связи, применение методов защиты биометрических шаблонов (ЗБШ) для модулей (систем) антиспуфинга является новой и нетривиальной исследовательской задачей [38].

Основной целью методов ЗБШ является защита биометрических шаблонов (уникальных физиологических характеристик человека, представленных в виде вектора) от несанкционированного доступа, подделки и компрометации. Современные исследователи [34] среди прочих выделяют три группы методов защиты биометрических шаблонов: отменяемая биометрия [35], биометрические криптосистемы [36] и гомоморфное шифрование [37]. Несмотря на характерные для каждой из групп достоинства, биометрические криптосистемы демонстрируют ряд существенных преимуществ перед методами отменяемой биометрии и гомоморфного шифрования. Биометрические криптосистемы используют криптографические ключи, связанные с биометрическими данными, что значительно снижает уязвимость системы к атакам. В случае утечки данных, криптографические

ключи можно изменить без необходимости повторного сбора биометрических данных. В отличие от отменяемой биометрии, где трансформации могут быть обратимыми и, следовательно, потенциально уязвимыми к атакам, биометрические криптосистемы обеспечивают более высокий уровень защиты за счёт сложности восстановления исходных данных из зашифрованных шаблонов. В сравнении с гомоморфным шифрованием, биометрические криптосистемы могут предложить более простую и менее ресурсоёмкую защиту, так как гомоморфное шифрование требует значительных вычислительных мощностей для обработки зашифрованных данных.

Одним из вариантов биометрических криптосистем, сочетающих в себе достоинства отменяемой биометрии и классических криптографических методов, являются схемы, обладающие возможностью связывания внешнего криптографического ключа с биометрическим образом [34]. К таким схемам относятся нейросетевые преобразователи «биометрия-код» (НПБК) [21], исключающие недостатки реализаций нечетких экстракторов [22] и обладающие возможностью анализа биометрических данных наравне с алгоритмами искусственного интеллекта. Основная цель НПБК — обеспечить надежное связывание биометрических данных с криптографическими ключами, которые затем используются для аутентификации и идентификации пользователей. Самой простой реализацией НПБК является преобразователь, обученный в соответствии с ГОСТ Р 52633.5-2011 [14], так называемый, классический НПБК. За счет своей уникальной структуры (широкой нейронной сети) такой НПБК способен обучаться на малых выборках биометрических образов, обогащая их в процессе обучения. Результатом обучения становятся высокие показатели точности распознавания биометрических образов и надежности процедуры аутентификации. Кроме того, согласно последним исследованиям [50], обозначенные показатели можно заметно улучшить за счет работы НПБК совместно с глубокими нейронными сетями, используемыми в качестве экстракторов признаков. Однако вопросы применимости таких реализаций для защиты альтернативных модулей систем биометрической аутентификации по лицу по-прежнему остаются не изученными.

Обзор существующих методов противодействия лицевым спуфинг атакам на основе глубокого обучения

Спуфинг атаки представляют собой вариант обмана биометрических систем аутентификации, при котором поддельные биометрические данные выдаются за подлинные с целью получения несанкционированного доступа. В контексте систем распознавания лиц, спуфинг атаки могут включать использование фотографий (распечатанных или демонстрируемых с цифровых устройств), видеозаписей, 3D-масок (распечатанных, силиконовых и т.д.) и других искусственных репрезентаций лица целевого пользователя. Для противодействия спуфинг атакам в системах распознавания лиц применяются различные методы, направленные на проверку подлинности представленных биометрических данных. Часто такие методы объединяют под общим термином liveness detection (методы определения живости), так как их главной задачей является ответ на вопрос «это живой (реальный) человек?». Отличать живые лица от подделок сегодня позволяют такие методы, как анализ текстуры кожи и изображений [29], наблюдение за микродвижениями [30, 31] (моргание или поворот головы), а также использование камер со специальными сенсорами [32].

Однако с появлением методов глубокого обучения (МГО) задача определения живого присутствия перестала быть трудоемким процессом, сопряжённым с ручным извлечением признаков поддельных изображений, характерных для традиционных методов liveness detection (LBP [5], SIFT [6], SURF [7], DoG [8], HOG [9] и другие). Теперь глубокие нейронные сети (ГНС) могут самостоятельно обнаруживать релевантные признаки, включая текстурные и контекстуальные особенности изображений настоящих и поддельных лиц. Кроме того, ГНС могут быть дополнительно улучшены методами регуляризации и предобучением на больших наборах данных, что увеличивает их способность к обобщению и устойчивости к различным типам атак.

Уже в 2014 году, авторы исследования [25] предложили первое полноценное решение систем определения живого присутствия на основе глубокого обучения с использованием 8-слойной неглубокой CNN для извлечения признаков. После этого до-

статочно часто стали появляться работы [26-28], в которых использовались уже предобученные глубокие модели (например, VGG16 или ResNet18), настраиваемые специально для задач определения живого присутствия с помощью трансферного обучения. Однако для того, чтобы оценить все разнообразие методов определения живого присутствия с помощью глубокого обучения, введем следующую классификацию таких методов [23]:

1. Гибридные методы: извлечение признаков с помощью классических методов с последующим применением глубоких моделей.

2. Традиционные методы обучения с учителем: так называемые end-to-end решения. Определение живого присутствия осуществляется путем применения методов глубокого обучения, чаще всего одной глубокой нейронной сети.

3. Методы, направленные на повышение обобщающей способности моделей глубокого обучения: подразумевают генерализацию моделей в отношении новых условий работы модели (освещение, качество изображения и др.) или новых типов атак.

4. Методы на основе дополнительной информации: используют специальные сенсоры или дополнительные модели для получения информации о входном изображении в иных диапазонах (инфракрасное излучение, тепловое излучение) или измерениях (карты глубины).

Наиболее простым вариантом обнаружения спуфинг атак с помощью метдов глубокого обучения является подход, при котором сначала извлекаются признаки из входных данных лица с помощью традиционных методов, а затем используется глубокое обучение для их семантического представления. Так, например, в работе [24] авторы используют LBP в качестве локальных дескрипторов, а затем работают с ними с помощью случайного леса. Стоит отметить, что исследователи не применяют CHC и при этом демонстрируют достаточно высокую эффективность на примере эталонного набора данных REPLAY-ATTACK [10]. Однако ключевым недостатком гибридных методов остается вся та же необходимость предварительного ручного извлечения признаков, свойственная традиционным подходам, а значит увеличение времени и ресурсов для обучения и настройки системы.

В большинстве работ, посвященных определению живого присутствия с использованием традиционных методов глубокого обучения, решение проблемы сводится к задаче бинарной классификации [39]. В таком случае модель обучается дифференцировать входные изображения по принципу «фальшивое лицо» (класс 1) и «реальное лицо» (класс 2). Такой подход является одним из наиболее простых и эффективных с точки зрения сходимости модели, однако не лишен недостатков: бинарная модель плохо поддается интерпретации, а выученные ею признаки сложно понять и использовать для улучшения производительности. Однако бинарная классификация способна демонстрировать довольно высокие результаты в случае учета временных характеристик распознавания (потока образов), что было продемонстрировано в работе [20] на примере модели FasTCo.

Отчасти с указанной проблемой способны справляться методы на основе попиксельного контроля (pixel-wise supervision) [4, 17, 19, 40] – подхода к обучению глубоких моделей, результатом которого становятся попиксельно маркированные изображения обучающего набора данных. Маркировка может осуществляться с целью получения карт псевдоглубины (pseudo depth labels) [17, 44], карт отражений (reflection maps) [45], карт бинарных масок (binary mask label) [46] или карт 3D облака точек (3D point cloud map) [47]. Идея получения таких различных карт при попиксельном контроле основывается на предположении о том, что подавляющее большинство спуфинг атак основаны на предъявлении системе двумерных изображений (распечаток или экранов устройств), в отличие от которых реальное лицо является объемным. Очевидно, что такой подход оказывается ограниченным с точки зрения работы с 3D атаками (масками), в связи с чем извлечение карт иного рода, например, карт бинарных масок [19], являются более предпочтительными для получения робастных моделей.

Одной из первых реализаций подхода стала архитектура DepthNet [2], до сих пор активно использующаяся для практических целей и работящая по принципу извлечения карт глубины (псевдоглубины) из входных изображений. Замена стандартных сверточных слоев DepthNet на специально разработанные для задачи антиспуфинга слои центральной разностной свертки (central

difference convolution (CDC)) позволили авторам работы [17] разработать новую архитектуру CDCN, демонстрирующую более высокие показатели точности распознавания спуфинг атак на протоколе №1 набора данных OULU-NPU [48] (ACER = 1.3%).

Однако высокие показатели метода попиксельного обучения на небольших наборах данных, вроде OULU-NPU [49] или SiW-M [49], трудно считать показательными, в особенности после появления одного из самых масштабных датасетов для задачи антиспуфинга – CelebA-Spoof [4]. Авторы коллекции утверждают, что одной из ключевых проблем развития моделей для обнаружения спуфинга являются наборы данных, плохо отражающие разнообразие спуфинг атак и их модификаций. С целью исправления ситуации был собран крупномасштабный набор данных и произведена качественная аннотация биометрических образов и атак. Дополнительно в работе [4] предложена одна из наиболее популярных моделей для задачи антиспуфинга AENet, за счет попиксельного обучения и механизмов внимания достигающая точности 99,6% на том же наборе данных.

Методы, направленные на повышение обобщающей способности моделей глубокого обучения, являются одними из самых перспективных среди современных подходов борьбы со спуфинг атаками. Актуальность направления связана с тем, что ГНС в принципе обладают слабой обобщающей способностью и часто «переучиваются» на специально подобранных наборах данных. Более того, согласно последним исследованиям [23], значительная часть работ по-прежнему опирается на небольшой пул устаревших наборов данных, которые трудно считать репрезентативными и пригодными для обучения моделей, предназначенных для работы в реальных условиях. В этом смысле показательными являются исследования, основанные на методах обучения с нулевым выстрелом (zero-shot learning) или с несколькими выстрелами (few-shot learning) [41], а также на методах обнаружения аномалий, в частности одноклассовой классификации. Так, например, в работе [15] авторы используют одноклассовый классификатор на основе многоканальной сверточной нейронной сети. Применение одноклассовой константной потери (one-class constative loss) для обучения классификатора позволило добиться однозначного разделения спуфинг образов и реальных изображений в пространстве векторных представлений. Такой подход обеспечивает возможность определения новых атак в реальных условиях. Аналогичного эффекта добиваются исследователи в работе [16], в качестве функции потерь использующие Hypersphere Loss Function. Применение функции позволяет добиться специфического распределения реальных образов в гиперсфере радиусом r и однозначно определять остальные образы как поддельные.

Наконец рассмотрим мультимодальный подход [42], применяющийся в настоящем исследовании и основанный на использовании дополнительной информации для обучения глубоких моделей. Основным достоинством подхода является возможность объединения информации из разных модальностей, что позволяет компенсировать недостатки каждой из них и дополнительно использовать различные алгоритмы и модели для анализа каждой из модальностей. Это позволяет оптимально использовать специализированные алгоритмы для обработки конкретного типа данных, улучшая общую производительность и точность системы. Кроме того, разные методы спуфинга могут быть направлены на уязвимости одной модальности, но крайне трудно обмануть систему, способную анализировать несколько модальностей одновременно.

Наиболее простым вариантом реализации подхода является использование специальных сенсоров для получения изображений в альтернативных диапазонах и обучение ГНС на полученных изображениях совместно со стандартными (RGB) [43]. Очевидны недостатки такого решения, связанные с необходимостью использования дополнительного оборудования. В связи с этим, дополнительной информацией для входов нейронной сети могут служить рассмотренные ранее карты, получаемые путем попиксельного контроля, например, карты псевдоглубины [3]. В данном случае неспособность карт глубины работать с 3D атаками частично компенсируется за счет дополнительного входа с классическим RGB изображением. Основной задачей для данного направления, в таком случае, становится поиск эффективных способов слияния информации от нескольких источников.

Из приведенного анализа видно, что наиболее перспективными методами противодействия спуфинг атакам в биометрических

системах аутентификации, в частности в системах распознавания лиц, являются методы на основе мультимодальных подходов и попиксельного контроля. Мультимодальные подходы позволяют объединять информацию из различных источников и биометрических модальностей, что компенсирует недостатки каждой из них и повышает общую точность и устойчивость системы к атакам. Методы попиксельного контроля, такие как карты псевдоглубины и отражений, обеспечивают детализированное представление изображений, что позволяет более эффективно различать реальные лица и подделки. Комбинирование этих методов может привести к созданию более робастных систем, способных противостоять различным видам атак, улучшая их производительность и надежность в реальных условиях.

Набор данных для экспериментальных исследований

Современные исследования в области лицевой биометрической аутентификации активно используют открытые наборы данных изображений и видеозаписей лиц для разработки и оценки алгоритмов определения подлинности лица. Специализированные наборы данных, такие как Replay-Attack [10], MSU-MFSD [11] и CelebA-Spoof [4], сфокусированы на задачах определения подлинности лиц и детекции спуфинг атак. Эти наборы данных включают изображения и видеозаписи различных атак, таких как использование фотографий, видео и 3D-масок. Наличие подобных специализированных наборов данных способствует возможности проведения экспериментальных оценок моделей глубокого обучения, направленных на выявление широкого спектра спуфинг атак.

Для проведения экспериментальных исследований, описанных ниже, за основу был взят один из немногих открытых наборов данных – CelebA-Spoof [4]. Набор данных CelebA-Spoof представляет собой обширную и тщательно аннотированную коллекцию изображений, специально созданную для задач анти-спуфинга в системах распознавания лиц. Датасет включает 625537 изображений 10177 субъектов, которые охватывают широкий спектр реальных лиц и различных типов атак, таких как атаки с использованием распечатанных изображений, атаки с использованием воспроизведения видео с различных устройств и распечатанные 3D маски. Каждое изображение в датасете сопровождается подробными аннотациями, включающими метки спуфинга (истинное или поддельное лицо), тип атаки (например, использование фотографий, видео или масок) и другие релевантные атрибуты. Изображения в датасете сделаны в различных условиях освещения, с разными позами и выражениями лиц.

Метрики оценки

В качестве основной метрики оценки работы предложенного решения используется точность распознавания (доля правильно классифицированных примеров относительно общего числа примеров). Для расчета указанной метрики использовалось следующее выражение:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

где TP (True Positives) — количество истинно положительных предсказаний, TN (True Negatives) — количество истинно отрицательных предсказаний, FP (False Positives) — количество ложно положительных предсказаний, FN (False Negatives) — количество ложно отрицательных предсказаний (когда модель ошибочно предсказала отрицательный класс).

В качестве дополнительной метрики высчитывается значение ACER (Average Classification Error Rate), являющееся общепринятой метрикой, использующейся для оценки общей производительности систем противодействия атакам на биометрическое предъявление и учитывающая ключевые ошибки – количество ложных принятий (APCER (Attack Presentation Classification Error Rate)) и ложных отказов (BPCER (Bona Fide Presentation Classification Error Rate)) – влияющие на надёжность системы.

$$ACER = \frac{APCER + BPCER}{2}$$

где $APCER = \frac{FP}{FP+TN}$, $BPCER = \frac{FN}{FN+TP}$.

Совместное использование указанных метрик обеспечивает всестороннюю оценку системы: высокое значение Accuracy указывает на общую эффективность классификации, тогда как низкое значение ACER свидетельствует о способности системы надежно различать реальные и поддельные образы.

Модуль обнаружения спуфинг атак для систем биометрической аутентификации по лицу

Для решения задачи обнаружения атак на биометрическое предъявление, при которой невозможен несанкционированный доступ к биометрическим данным пользователей, в настоящей работе предложен модуль обнаружения спуфинг атак для систем биометрической аутентификации по лицу с использованием нейросетевого преобразователя «биометрия-код».

Одним из способов нивелирования недостатков применения глубоких нейронных сетей для обнаружения спуфинг атак является логика разделения блока векторного представления образов и блока защиты биометрического шаблона (блока классификации) (рис. 1). Если в качестве последнего для классификации образов использовать широкие нейронные сети, способные обучаться автоматически (без применения градиентного спуска), то можно избежать переобучения всей сети на специальном наборе данных и несколько повысить обобщающую способность модели за счет классификатора, обученного общему представлению реальных и поддельных изображений. В таком случае, спектр атак, представленный в наборе данных CelebA-Spoof, не оказывает решающего значения при обучении и тестировании предложенных моделей.

Как отмечалось ранее, в качестве описанной широкой нейронной сети может выступать нейросетевой преобразователь «биометрия-код», позволяющий достигать сразу двух ключевых целей в разработке модуля обнаружения спуфинг атак: противодействие угрозам компрометации биометрических образов пользователей, проходящим через систему обнаружения, а также высокая точность классификации реальных и поддельных изображений лиц.

Дополнительным достоинством всего предложенного решения (модуля), является отсутствие необходимости полного повторного переобучения модели глубокого обучения (блока векторного представления), предшествующей НПБК, в случае компрометации биометрических образов лиц на этапе обнаружения спуфинг атак.

В рамках экспериментальной реализации модуля обнаружения спуфинг атак в качестве блока векторного представления была обучена глубокая нейронная сеть, основанная на архитектуре FeatherNet [3]. FeatherNet — это легковесная архитектура сверточной нейронной сети, разработанная для задачи обнаружения спуфинг атак в системах распознавания лиц. В основе архитектуры лежат идеи минимизации вычислительных затрат и параметров модели без потери точности.

Одной из ключевых особенностей оригинальной архитектуры FeatherNet, предложенной в работе [3], является замена широко применяемого для снижения размерности слоя глобального усредняющего пулинга (Global Average Pooling (GAP)) на так называемый модуль потоковой передачи данных (Streaming Module), основанный на глубинной свертке (depthwise convolution) с шагом > 1. С помощью такой замены удается избегать негативных последствий применения GAP для задач распознавания лиц, связанных с усреднением всех значений карт признаков вне зависимости от степени их «важности» для конкретного примера. Кроме того, FeatherNet является примером мультимодальных архитектур, прини-

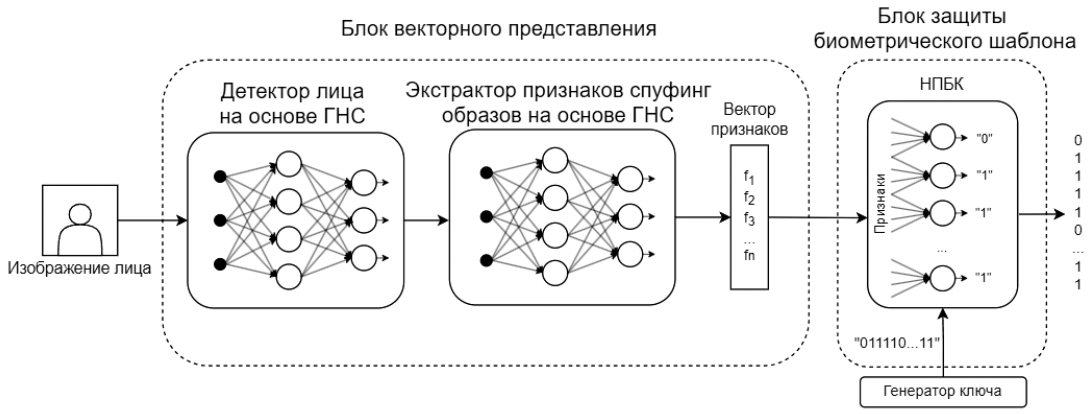

Рис. 1. Схема модуля обнаружения спуфинг атак для систем биометрической аутентификации по лицу

Таблица 1.

Архитектура модифицированной сверточной нейронной сети для извлечения признаков на основе FeatherNet

Назначение	Слой	Описание	Размерность выходного вектора
Усреднение мульти-модальных входных значений	conv1	Блок с Conv2d_cdcn (3 входных канала, 32 выходных, ядро 3x3, шаг 2), BatchNorm2d и Hardswish активацией.	32
	avgpool	AdaptiveAvgPool2d, выполняет адаптивное усреднение по размерам входного изображения для слоя layer1.	32
Извлечение признаков	layer1	Блоки, содержащие InvertedResidual блоки, включая Downsample, Conv2d, BatchNorm2d, Hardswish и SELayer (для некоторых блоков).	16
	layer2		32
	layer3		64
	layer4		96
Классификаторы («головы»)	fc_face	Блок с Dropout (вероятность отсева 0.15), BatchNorm1d, Hardswish и Linear (входных 96, выходных 40).	40
	fc_attack	Блок с Dropout (вероятность отсева 0.15), BatchNorm1d, Hardswish и Linear (входных 96, выходных 11).	11
	fc_light	Блок с Dropout (вероятность отсева 0.15), BatchNorm1d, Hardswish и Linear (входных 96, выходных 5).	5
	fc_live	Блок с Dropout (вероятность отсева 0.15), BatchNorm1d, Hardswish и AngleSimpleLinear (вычисление косинуса угла между векторами в пространстве признаков).	1
	depth	Блок с Conv2d (входных 96, выходных 1), Upsample (размер 14x14, метод 'bilinear') и Sigmoid активацией.	1

мающих на вход не только стандартные изображения (RGB), но и дополнительную информацию в виде карт глубины [44] и инфракрасных изображений лиц [42].

Для обучения описанной архитектуры и получения блока векторного представления (экстрактора признаков), способного дифференцировать признаки, полученные из изображений реальных и поддельных лиц, были произведены ряд модификаций в структуре сети FeatherNet (табл. 1). Целью модификаций является объединение преимуществ технологий попиксельного контроля (центральная разностная свертка) и мультимодальных архитектур для повышения эффективности работы итоговой модели в отношении разнообразия спуфинг атак.

В первую очередь, для обучения сети кроме стандартных RGB изображений использовались только карты глубины лиц, представленные во вспомогательном наборе данных CelebA-Spoof Depth Image [33]. Исключение изображений в инфракрасном диапазоне из процедуры обучения обосновано отсутствием таковых для экспериментального набора данных CelebA-Spoof.

В качестве дополнительных улучшений были произведены замены функций активаций ReLU (Rectified Linear Unit) на относительно новый тип нелинейности для обучения глубоких нейронных сетей HardSwish (h-swish), впервые представленный в рамках исследований архитектуры MobileNetV3 [18]:

$$hardswish(x) = x\frac{ReLU6(x+3)}{6}$$

где ReLU6 – это функция активации ReLU, ограниченная значением 6. HardSwish явля-

ется аппроксимацией swish, которая умножает входное значение x на сигмовидную функцию от этого же значения. Для современных архитектур, разработанных для мобильных и встраиваемых систем, использование функции активации HardSwish позволяет достигать высокого соотношения производительности и затрат.

Кроме того, для блока усреднения мультимодальных входных значений были изменены процедуры свертки, позволяющие извлекать пространственные признаки из входных данных: классическая 2D свертка (Conv2d) в настоящей работе заменяется на ее модификацию, предназначенную для улучшения способности модели извлекать локальные градиентные признаки в данных – central difference convolution (CDC) [17]. Центральная разностная свертка работает за счет добавления к процедуре стандартной свертки входного изображения с фильтром K дополнительной информации в виде разности каждого элемента окна свертки и центрального элемента этого окна, умноженной на соответствующий элемент фильтра. Добавление описанной модификации в первые слои усреднения позволяют повысить устойчивость модели к небольшим изменениям и искажениям в данных, что является критически важным для возможности распознавания максимально приближенных к реальности спуфинг атак.

Экспериментальная оценка предложенного решения

Для повышения обобщающей способности модели, а также предотвращения возможного переобучения была произведена аугментация данных тренировочного набора. Для этого входные изображения 7 видам дополнительных преобразований:

1. Добавление шума, имитирующего шум цифровых камер (ISOnoise), с определенным сдвигом цвета и интенсивностью. Применяется с вероятностью 5%.

2. Изменение яркости и контраста в заданных пределах. Применяется с вероятностью 12.5%.

3. Имитация движения, осуществляющее размытие, имитирующее движение, с ограничением размытия до 3 пикселей. Применяется с вероятностью 20%.

4. Имитация сжатия изображения, помогающее снижать его качество. Применяется с вероятностью 25%.

5. Случайное удаление фрагментов изображения (заполнение черным) в виде прямоугольных областей. Применяется с вероятностью 25%.

6. Добавление гауссовского шума. Применяется с вероятностью 20%.

7. Нормализация изображения с заданными средними значениями и стандартными отклонениями для каждого канала (RGB).

Применение указанных видов аугментации способно значительно повысить эффективность антиспуфинг систем распознавания лиц за счет увеличения разнообразия обучающих данных и повышения устойчивости моделей к различным искажениям. Добавление шума, изменение яркости и контраста, а также имитация движения и сжатия изображений способствуют адаптации моделей к вариативным условиям съемки и различным уровням качества изображений. Случайное удаление фрагментов изображения и добавление гауссовского шума увеличивают способность моделей справляться с частичными потерями информации и шумами. Нормализация изображений обеспечивает стабильность процесса обучения, выравнивая данные и ускоряя сходимость моделей.

Обучение модифицированной архитектуры FeatherNet осуществлялось на полном наборе данных CelebA-Spoof в течение 15 эпох с помощью функции потерь Cross Entropy для «головы» fc_live, осуществляющей бинарную классификацию реальных и поддельных изображений (рис. 2). Предварительно обучающие данные были разделены на тренировочную и тестовую (валидационную) выборки, каждая из которых случайным образом подвергалась аугментации в соответствии с описанными выше преобразованиями.

На вариационной выборке максимальное значение точности работы сети составило только 92,7%, несмотря на то, что значения точности на тренировочных данных превышают 97%. Для дальнейшего использования обученной сети в качестве экстрактора признаков, слои классификатора «замораживались», а работа осуществлялась с 96-мерным вектором признаков на выходе предшествующего классификатору слоя.

В качестве детектора лиц на входных изображениях использовалась предобученная на крупномасштабном наборе данных WIDERFACE [12] модель RetinaFace [13]. Модель разработана для одностадийного обнаружения лиц с высокой точностью и использует стратегии многозадачного обучения.

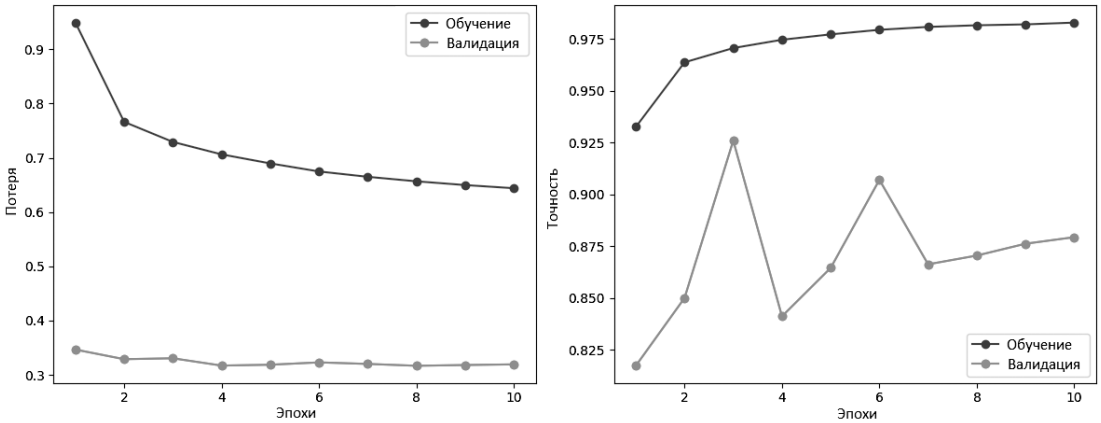

Рис. 2. Результаты обучения модифицированной архитектуры FeatherNet

В качестве классификатора реальных и поддельных входных образов лиц реализован нейросетевой преобразователь биометрия-код, процедура обучения которого представлена в ГОСТ Р 52633.5-2011. Адаптация логики работы НПБК под задачу бинарной классификации осуществляется путем построения преобразователя для реальных изображений лиц (класс C_1). В таком случае, поддельные изображения (класс C_2) расцениваются НПБК как «Чужие», а случайный код на выходе свидетельствует о решении в пользу класса C_2. Принадлежность классу C_i, $i = 1, 2$, оценивается исходя из получаемого на выходе НПБК бинарного кода (ключ K_2 в случае класса C_1). Особенностью построения НПБК в качестве классификатора является допущение о возможности дублировании входов нейрона в связи с отсутствием необходимости сокрытия структуры преобразователя. Для оценки качества осуществляемой бинарной классификации i-ого входного образа применяется следующее правило:

$$\begin{cases} (\bar{a}_i \in C_1 \ \wedge \ h_i < \ threshold) \rightarrow \ TP \\ (\bar{a}_i \in C_1 \ \wedge \ h_i > \ threshold) \rightarrow \ FN \\ (\bar{a}_i \in C_2 \ \wedge \ h_i < \ threshold) \rightarrow \ FP \\ (\bar{a}_i \in C_2 \ \wedge \ h_i > \ threshold) \rightarrow \ TN \end{cases}$$

где h_i – i-ое значение расстояния Хэмминга между ожидаемым кодом и выходом НПБК, $threshold$ – порог, определяющий допустимое количество ошибок в коде i-ого входного образа для корректного отнесения его к одной из групп классифицированных образов: TP – True Positive, FN – False Negative, FP – False Positive или TN – True Negative. Полученный классификатор не требует итерационного обучения (осуществляется автоматически) и большого числа обучающих примеров.

Итоговый эксперимент по оценке точности работы нейросетевого преобразователя «биометрия-код» в качестве классификатора (рис. 3)

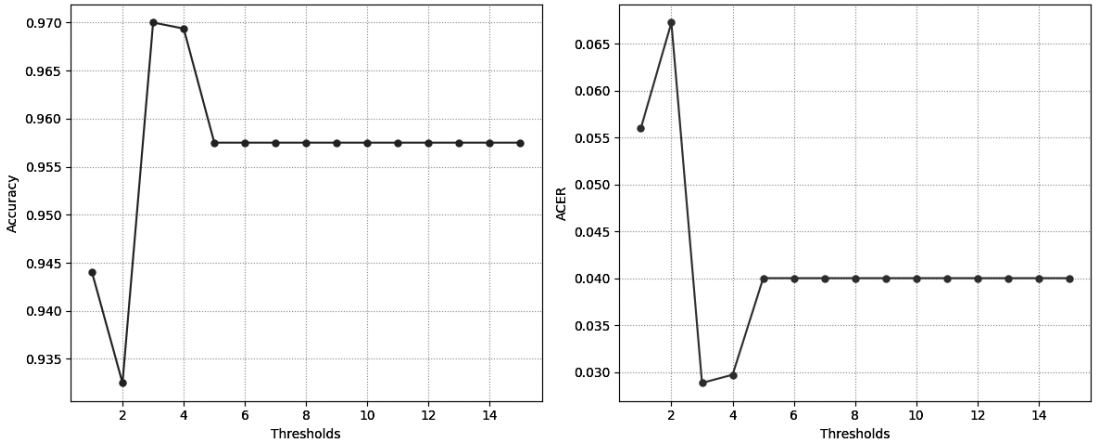

Рис. 3. Точность (accuracy) работы НПБК при разных значениях threshold

проводился с помощью специально подготовленных для этой задачи выборок (тренировочной и тестовой) из датасета CelebA-Spoof. В случае обучающей выборки случайным образом из классов (субъектов) основного набора данных выбирались 100 изображений реальных лиц, представляющих собой класс C_1, и 100 изображений поддельных лиц, представляющих собой разнообразные атаки исходного набора данных CelebA-Spoof (класс C_2). Аналогичная процедура была произведена для получения тестовой выборки, однако полученные 200 изображений дополнительно случайным образом перемешивались.

Анализ результатов работы НПБК при различных значениях threshold (рис. 3) продемонстрировал, что оптимальное значение порога, при котором достигается наивысшая точность (97.0%) и минимальная среднеклассовая ошибка (ACER = 2.9%) составляет threshold = 3. Это значение обеспечивает наилучшую производительность модели в задаче классификации реальных и поддельных биометрических данных. При других значениях порога наблюдается снижение точности до уровня около 95.5% и увеличение ACER, особенно при threshold = 2, где ошибка превышает 6.5%, что свидетельствует о значительном ухудшении качества классификации. Таким образом, пороговое значение threshold = 3 обеспечивает наилучшее соотношение между точностью распознавания и надежностью работы модуля, а также является наиболее оптимальным для задачи классификации реальных и поддельных биометрических лицевых данных.

Сравнение предложенного решения с существующими аналогами на основе глубокого обучения

Результаты сравнения предложенного решения с классическими моделями глубокого обучения, используемыми для задачи обнаружения спуфинг атак, представлены в таблице 2. Несмотря на то, что в части исследований [15, 17, 19, 20] авторы не предоставляют информации о результирующей точности работы сети, оценку моделей можно производить исходя из метрики ACER.

Из таблицы видно, что разработанный модуль (FeatherNet + НПБК) демонстрирует точность 97,2%, что уступает только модели AENet (99,6%), однако по-прежнему показывает высокий уровень распознавания реальных и поддельных лиц. В свою очередь, значение ACER для предложенного решения составляет 2,9%, что ниже, чем у AENet (3,09%) и значительно ниже по сравнению с моделями DeepPixBiS (5,97%) и CDCN++ (1,3%). Такие показатели ACER свидетельствуют о достаточно высокой надежности модели при условии обеспечения безопасного режима работы антиспуфинг модуля.

Показатели ACER для моделей MCCNN (BCE+OCCL)-GMM и FasTCo на наборе данных SiW-M демонстрируют большую вариативность (14.9 ± 7.8% и 10.1 ± 5.6% соответственно), что указывает на их меньшую стабильность по сравнению с FeatherNet + НПБК, а также на тот факт, что глубокие нейронные сети для обнаружения спуфинг атак, зачастую, проучиваются на специализированных наборах данных и показатели их эффективности, полученные в результате обучения, мо-

Таблица 2.

Сравнительные результаты работы предложенного решения с моделями глубокого обучения для обнаружения спуфинг атак

№	Модель	Набор данных	Accuracy	ACER
1	DeepPixBiS [19]	OULU-NPU (p.2)	-	5.97%
2	CDCN++ [17]	OULU-NPU (p.1)	-	1.3%
3	AENet [4]	CelebA-Spoof Dataset	99,60%	3.09%
4	MCCNN (BCE+OCCL)-GMM [15]	SiW-M	-	14.9 ± 7.8%
5	FasTCo [20]	SiW-M	-	10.1 ± 5.6%
7	FeatherNet + НПБК	CelebA-Spoof Dataset	97,20%	2,90%

гут существенно отличатся от значений, полученных в реальных условиях функционирования системы.

Несмотря на сравнительно невысокие показатели точности работы НПБК в качестве классификатора, сохраняется ключевое преимущество предложенного решения: безопасная реализация биометрической аутентификации по лицу, при которой модуль обнаружения спуфинг атак перестает быть точкой потенциальных уязвимостей. По-видимому, применение методов защиты биометрических шаблонов лиц к решениям для обнаружения спуфинг атак может привести к снижению точности распознавания, но взамен значительно повышает безопасность процедуры аутентификации с интегрированными модулями антиспуфинга.

Заключение

Настоящая работа посвящена исследованию применимости нейросетевых преобразователей «биометрия-код» для защиты блоков обнаружения спуфинг атак в системах биометрической аутентификации по лицу. С этой целью разработан и предложен антиспуфинг модуль на основе НПБК и глубокой нейронной сети, осуществляющий бинарную классификацию входных образов лиц на реальные и поддельные изображения. За счет разделения блока векторного представления образов (глубокой нейронной сети) и блока принятия решения в виде классификатора на основе НПБК, повышается обобщающая способность предложенного решения по отношению к разнообразию реализации спуфинг атак и решается проблема несанкционированного доступа в системах биометрической аутентификации по лицу, осуществляемого через блоки обнаружения спуфинг атак. Лучшее значение точности работы модуля на наборе данных CelebA-Spoof составило 97,2% (ACER = 2,9%), что говорит о приемлемом уровне производительности решения при высоком уровне защищенности процедуры аутентификации.

Литература

1. ГОСТ Р 58624.1-2019. Информационные технологии. Биометрия. Обнаружение атаки на биометрическое предъявление. Часть 1. Структура – Введ. 01.06.2020. – М.: Стандарт информ, 2019. – 11 с. – (Система стандартов по информации, библиотечному и издательскому делу).

2. Y. Atoum, Y. Liu, A. Jourabloo, and X. Liu, "Face anti-spoofing using patch and depth-based CNNs," in Proc. IEEE Int. Joint Conf. Biometrics, 2017, pp. 319–328.

3. Zhang P. et al. FeatherNets: Convolutional neural networks as light as feather for face anti-spoofing //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. – 2019. – С. 0-0.

4. Zhang Y. et al. Celeba-spoof: Large-scale face anti-spoofing dataset with rich annotations // Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16. – Springer International Publishing, 2020. – С. 70-85.

5. Das, D., & Chakraborty, S. Face liveness detection based on frequency and micro-texture analysis // 2014 International Conference on Advances in Engineering & Technology Research (ICAETR), 2014.

6. Patel, K., Han, H., & Jain, A. K. Secure Face Unlock: Spoof Detection on Smartphones // IEEE Transactions on Information Forensics and Security, 11(10), 2016.

7. Boulkenafet, Z., Komulainen, J., & Hadid, A. Face Anti-Spoofing using Speeded-Up Robust Features and Fisher Vector Encoding // IEEE Signal Processing Letters, 2016.

8. Tan, X., Li, Y., Liu, J., & Jiang, L. Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model // Lecture Notes in Computer Science, 2010.

9. Komulainen, J., Hadid, A., & Pietikainen, M. Context based face anti-spoofing // 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013.

10. Chingovska I., Anjos A., Marcel S. On the effectiveness of local binary patterns in face anti-spoofing //2012 BIOSIG-proceedings of the international conference of biometrics special interest group (BIOSIG). – IEEE, 2012. – С. 1-7.

11. Wen D., Han H., Jain A. K. Face spoof detection with image distortion analysis //IEEE Transactions on Information Forensics and Security. – 2015. – Т. 10. – №. 4. – С. 746-761.

12. Yang S. et al. Wider face: A face detection benchmark //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 5525-5533.

13. Deng J. et al. Retinaface: Single-stage dense face localisation in the wild //arXiv preprint arXiv:1905.00641. – 2019.

14. ГОСТ Р 52633.5-2011. Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа – Введ. 01.04.2012. – М.: Стандарт информ, 2018. – 13 с. – (Система стандартов по информации, библиотечному и издательскому делу).

15. George, Anjith, and Sébastien Marcel. "Learning one class representations for face presentation attack detection using multi-channel convolutional neural networks." IEEE Transactions on Information Forensics and Security 16 (2020): 361-375.

16. Li, Zhi, et al. "Unseen face presentation attack detection with hypersphere loss." ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020.

17. Yu Z. et al. Searching central difference convolutional networks for face anti-spoofing //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2020. – С. 5295-5305.

18. Howard A. et al. Searching for mobilenetv3 //Proceedings of the IEEE/CVF international conference on computer vision. – 2019. – С. 1314-1324.

19. George A., Marcel S. Deep pixel-wise binary supervision for face presentation attack detection //2019 International Conference on Biometrics (ICB). – IEEE, 2019. – С. 1-8.

20. Xu X., Xiong Y., Xia W. On improving temporal consistency for online face liveness detection system //Proceedings of the IEEE/CVF International Conference on Computer Vision. – 2021. – С. 824-833.

21. B. S. Akhmetov, A. I. Ivanov, and Z. Alimseitova, 'Training of neural network biometry-code converters', News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences, vol. 1, pp. 61–68, Jan. 2018.

22. Sarkar, A., & Singh, B. K. (2020). A review on performance, security and various biometric template protection schemes for biometric authentication systems. Multimedia Tools and Applications

23. Zitong Yu, Yunxiao Qin, Xiaobai Li, Chenxu Zhao, Zhen Lei, Guoying Zhao. Deep Learning for Face Anti-Spoofing: A Survey // IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2022.

24. R. Cai and C. Chen. Learning deep forest with multi-scale local binary pattern features for face anti-spoofing // arXiv preprint, 2019.

25. J. Yang, Z. Lei, and S. Z. Li. Learn convolutional neural network for face anti-spoofing // arXiv preprint, 2014.

26. O. Lucena, A. Junior, V. Moia, R. Souza, E. Valle, and R. Lotufo. Transfer learning using convolutional neural networks for face anti-spoofing // ICIAR, 2017.

27. H. Chen, G. Hu, Z. Lei, Y. Chen, N. M. Robertson, and S. Z. Li. Attention-based two-stream convolutional networks for face spoofing detection // TIFS, 2019.

28. A. George and S. Marcel. On the effectiveness of vision transformers for zero-shot face anti-spoofing // arXiv preprint, 2020.

29. I Chingovska, A Anjos, Sébastien Marcel. On the Effectiveness of Local Binary Patterns in Face Anti-spoofing // IEEE International Conference of the Biometrics Special Interest Group (BIOSIG), 2012.

30. Anjos, A., Chakka, M. M., & Marcel, S. Motion-based counter-measures to photo attacks in face recognition // IET Biometrics, 3(3), 2014.

31. G. Pan, L. Sun, Z. Wu, and S. Lao. Eyeblink-based anti-spoofing in face recognition from a generic webcamera // ICCV, 2007.

32. 32. Mohamed, S., Ghoneim, A., & Youssif, A. Visible/Infrared face spoofing detection using texture descriptors // MATEC Web of Conferences, 2019.

33. 33. Kaggle. CelebA-Spoof Depth Image. Available at: https://www.kaggle.com/datasets/attentionlayer241/celeba-spoof-depth-image (accessed 1 August 2024).

34. Sarkar A., Singh B. K. A review on performance, security and various biometric template protection schemes for biometric authentication systems //Multimedia Tools and Applications. – 2020. – Т. 79. – №. 37. – С. 27721-27776.

35. Manisha, Kumar N. Cancelable biometrics: a comprehensive survey //Artificial Intelligence Review. – 2020. – Т. 53. – №. 5. – С. 3403-3446.

36. Kaur P., Kumar N., Singh M. Biometric cryptosystems: a comprehensive survey //Multimedia Tools and Applications. – 2023. – Т. 82. – №. 11. – С. 16635-16690.

37. Bassit A. et al. Hybrid biometric template protection: Resolving the agony of choice between bloom filters and homomorphic encryption //IET biometrics. – 2022. – Т. 11. – №. 5. – С. 430-444.

38. Song, Baogang, et al. "An Investigation of the Effectiveness of Template Protection Methods on Protecting Privacy During Iris Spoof Detection." Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data. Singapore: Springer Nature Singapore, 2023.

39. Yang, Jianwei, Zhen Lei, and Stan Z. Li. "Learn convolutional neural network for face anti-spoofing." arXiv preprint arXiv:1408.5601 (2014).

40. Y. Qin, Z. Yu, L. Yan, Z.Wang, C. Zhao, and Z. Lei, "Meta-teacher for face anti-spoofing," TPAMI, 2021.

41. Y. Qin, C. Zhao, X. Zhu, Z. Wang, Z. Yu, T. Fu, F. Zhou, J. Shi, and Z. Lei, "Learning meta model for zero- and few-shot face anti-spoofing," in AAAI, 2020.

42. A. George and S. Marcel, "Cross modal focal loss for rgbd face anti-spoofing," in CVPR, 2021.

43. A. Liu, Z. Tan, J. Wan, Y. Liang, Z. Lei, G. Guo, and S. Z. Li, "Face anti-spoofing via adversarial cross-modality translation," TIFS, 2021.

44. D. Peng, J. Xiao, R. Zhu, and G. Gao, "Ts-fen: Probing feature selection strategy for face anti-spoofing," in ICASS. IEEE, 2020.

45. Z. Yu, X. Li, X. Niu, J. Shi, and G. Zhao, "Face anti-spoofing with human material perception," in ECCV, 2020.

46. A. George and S. Marcel, "Deep pixel-wise binary supervision for face presentation attack detection," in ICB, no. CONF, 2019.

47. X. Li, J. Wan, Y. Jin, A. Liu, G. Guo, and S. Z. Li, "3dpc-net: 3d point cloud network for face anti-spoofing," 2020.

48. Boulkenafet Z. et al. OULU-NPU: A mobile face presentation attack database with real-world variations //2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017). – IEEE, 2017. – С. 612-618.

49. Liu Y., Jourabloo A., Liu X. Learning deep models for face anti-spoofing: Binary or auxiliary supervision //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – С. 389-398.

50. Sulavko A. Biometric-based key generation and user authentication using acoustic characteristics of the outer ear and a network of correlation neurons //Sensors. – 2022. – Т. 22. – №. 23. – С. 9551.

References

1. ГОСТ Р 58624.1-2019. Informatsionnye tekhnologii. Biometriya. Obnaruzhenie ataki na biometricheskoe pred"yavlenie. Chast' 1. Struktura – Vved. 01.06.2020. – M.: Standart inform, 2019. – 11 s. – (Sistema standartov po informatsii, bibliotechnomu i izdatel'skomu delu).

2. Atoum Y., Liu Y., Jourabloo A., Liu X. Face Anti-Spoofing Using Patch and Depth-Based CNNs. In: Proc. IEEE Int. Joint Conf. Biometrics, 2017, pp. 319–328.

3. Zhang P. et al. FeatherNets: Convolutional Neural Networks as Light as Feather for Face Anti-Spoofing. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019, pp. 0-0.

4. Zhang Y. et al. Celeba-Spoof: Large-Scale Face Anti-Spoofing Dataset with Rich Annotations. In: Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16, Springer International Publishing, 2020, pp. 70-85.

5. Das D., Chakraborty S. Face Liveness Detection Based on Frequency and Micro-Texture Analysis. In: 2014 International Conference on Advances in Engineering & Technology Research (ICAETR), 2014.

6. Patel K., Han H., Jain A.K. Secure Face Unlock: Spoof Detection on Smartphones. In: IEEE Transactions on Information Forensics and Security, 2016, vol. 11, no. 10.

7. Boulkenafet Z., Komulainen J., Hadid A. Face Anti-Spoofing Using Speeded-Up Robust Features and Fisher Vector Encoding. In: IEEE Signal Processing Letters, 2016.

8. Tan X., Li Y., Liu J., Jiang L. Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model. In: Lecture Notes in Computer Science, 2010.

9. Komulainen J., Hadid A., Pietikainen M. Context Based Face Anti-Spoofing. In: 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013.

10. Chingovska I., Anjos A., Marcel S. On the Effectiveness of Local Binary Patterns in Face Anti-Spoofing. In: 2012 BIOSIG-proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG), IEEE, 2012, pp. 1-7.

11. Wen D., Han H., Jain A.K. Face Spoof Detection with Image Distortion Analysis. In: IEEE Transactions on Information Forensics and Security, 2015, vol. 10, no. 4, pp. 746-761.

12. Yang S. et al. Wider Face: A Face Detection Benchmark. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 5525-5533.

13. Deng J. et al. RetinaFace: Single-Stage Dense Face Localisation in the Wild. In: arXiv preprint arXiv:1905.00641, 2019.

14. ГОСТ Р 52633.5-2011. Zashchita informatsii. Tekhnika zashchity informatsii. Avtomaticheskoe obuchenie neiurosetevykh preobrazovatelei biometriya-kod dostupa – Vved. 01.04.2012. – M.: Standart inform, 2018. – 13 s. – (Sistema standartov po informatsii, bibliotechnomu i izdatel'skomu delu).

15. George A., Marcel S. Learning One Class Representations for Face Presentation Attack Detection Using Multi-Channel Convolutional Neural Networks. In: IEEE Transactions on Information Forensics and Security, 2020, vol. 16, pp. 361-375.

16. Li Z. et al. Unseen Face Presentation Attack Detection with Hypersphere Loss. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020.

17. Yu Z. et al. Searching Central Difference Convolutional Networks for Face Anti-Spoofing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5295-5305.

18. Howard A. et al. Searching for MobileNetV3. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 1314-1324.

19. George A., Marcel S. Deep Pixel-Wise Binary Supervision for Face Presentation Attack Detection. In: 2019 International Conference on Biometrics (ICB), IEEE, 2019, pp. 1-8.

20. Xu X., Xiong Y., Xia W. On Improving Temporal Consistency for Online Face Liveness Detection System. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 824-833.

21. Akhmetov B.S., Ivanov A.I., Alimseitova Z. Training of Neural Network Biometry-Code Converters. In: News of the National Academy of Sciences of the Republic of Kazakhstan, Series of Geology and Technical Sciences, 2018, vol. 1, pp. 61–68.

22. Sarkar A., Singh B.K. A Review on Performance, Security and Various Biometric Template Protection Schemes for Biometric Authentication Systems. In: Multimedia Tools and Applications, 2020.

23. Yu Z., Qin Y., Li X., Zhao C., Lei Z., Zhao G. Deep Learning for Face Anti-Spoofing: A Survey. In: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2022.

24. Cai R., Chen C. Learning Deep Forest with Multi-Scale Local Binary Pattern Features for Face Anti-Spoofing. In: arXiv preprint, 2019.

25. Yang J., Lei Z., Li S.Z. Learn Convolutional Neural Network for Face Anti-Spoofing. In: arXiv preprint, 2014.

26. Lucena O., Junior A., Moia V., Souza R., Valle E., Lotufo R. Transfer Learning Using Convolutional Neural Networks for Face Anti-Spoofing. In: ICIAR, 2017.

27. Chen H., Hu G., Lei Z., Chen Y., Robertson N.M., Li S.Z. Attention-Based Two-Stream Convolutional Networks for Face Spoofing Detection. In: TIFS, 2019.

28. George A., Marcel S. On the Effectiveness of Vision Transformers for Zero-Shot Face Anti-Spoofing. In: arXiv preprint, 2020.

29. Chingovska I., Anjos A., Marcel S. On the Effectiveness of Local Binary Patterns in Face Anti-Spoofing. In: IEEE International Conference of the Biometrics Special Interest Group (BIOSIG), 2012.

30. Anjos A., Chakka M.M., Marcel S. Motion-Based Counter-Measures to Photo Attacks in Face Recognition. In: IET Biometrics, 2014, vol. 3, no. 3.

31. Pan G., Sun L., Wu Z., Lao S. Eyeblink-Based Anti-Spoofing in Face Recognition from a Generic Webcamera. In: ICCV, 2007.

32. Mohamed S., Ghoneim A., Youssif A. Visible/Infrared Face Spoofing Detection Using Texture Descriptors. In: MATEC Web of Conferences, 2019.

33. Kaggle. CelebA-Spoof Depth Image. Available at: https://www.kaggle.com/datasets/attentionlayer241/celeba-spoof-depth-image (accessed 1 August 2024).

34. Sarkar A., Singh B.K. A Review on Performance, Security and Various Biometric Template Protection Schemes for Biometric Authentication Systems. In: Multimedia Tools and Applications, 2020, vol. 79, no. 37, pp. 27721-27776.

35. Manisha, Kumar N. Cancelable Biometrics: A Comprehensive Survey. In: Artificial Intelligence Review, 2020, vol. 53, no. 5, pp. 3403-3446.

36. Kaur P., Kumar N., Singh M. Biometric Cryptosystems: A Comprehensive Survey. In: Multimedia Tools and Applications, 2023, vol. 82, no. 11, pp. 16635-16690.

37. Bassit A. et al. Hybrid Biometric Template Protection: Resolving the Agony of Choice between Bloom Filters and Homomorphic Encryption. In: IET Biometrics, 2022, vol. 11, no. 5, pp. 430-444.

38. Song B. et al. An Investigation of the Effectiveness of Template Protection Methods on Protecting Privacy During Iris Spoof Detection. In: Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data, Singapore: Springer Nature Singapore, 2023.

39. Yang J., Lei Z., Li S.Z. Learn Convolutional Neural Network for Face Anti-Spoofing. In: arXiv preprint arXiv:1408.5601, 2014.

40. Qin Y., Yu Z., Yan L., Wang Z., Zhao C., Lei Z. Meta-Teacher for Face Anti-Spoofing. In: TPAMI, 2021.

41. Qin Y. et al. Learning Meta Model for Zero-and Few-Shot Face Anti-Spoofing. In: AAAI, 2020.

42. George A., Marcel S. Cross Modal Focal Loss for RGBD Face Anti-Spoofing. In: CVPR, 2021.

43. Liu A. et al. Face Anti-Spoofing via Adversarial Cross-Modality Translation. In: TIFS, 2021.

44. Peng D. et al. TS-FEN: Probing Feature Selection Strategy for Face Anti-Spoofing. In: ICASS, IEEE, 2020.

45. Yu Z. et al. Face Anti-Spoofing with Human Material Perception. In: ECCV, 2020.

46. George A., Marcel S. Deep Pixel-Wise Binary Supervision for Face Presentation Attack Detection. In: ICB, 2019, no. CONF.

47. Li X. et al. 3DPC-Net: 3D Point Cloud Network for Face Anti-Spoofing, 2020.

48. Boulkenafet Z. et al. OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations. In: 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), IEEE, 2017, pp. 612-618.

49. Liu Y., Jourabloo A., Liu X. Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 389-398.

50. Sulavko A. Biometric-Based Key Generation and User Authentication Using Acoustic Characteristics of the Outer Ear and a Network of Correlation Neurons. In: Sensors, 2022, vol. 22, no. 23, p. 9551.

ПАНФИЛОВА Ирина Евгеньевна, аспирант федерального государственного бюджетного образовательного учреждения высшего образования «Самарский государственный технический университет». 443100, г. Самара, ул. Молодогвардейская, 244. E-mail: panfilova_2015@bk.ru

ЛОЖНИКОВ Павел Сергеевич, проректор по научной и инновационной деятельности федерального государственного автономного образовательного учреждения высшего образования «Омский государственный технический университет». 644050, г. Омск, пр. Мира, 11. E-mail: lozhnikov@gmail.com

PANFILOVA Irina Evgenevna, postgraduate student of the Federal State Budgetary Educational Institution of Higher Education "Samara State Technical University". 443100, Samara, Molodogvardeyskaya str. 244. E-mail: panfilova_2015@bk.ru

LOZHNIKOV Pavel Sergeevich, Vice-Rector for Research and Innovation Activity of the Federal State Autonomous Educational Institution of Higher Education "Omsk State Technical University". 11 Mira Ave., Omsk, 644050, Omsk. E-mail: lozhnikov@gmail.com

Материалы к публикации отправлять по адресу E-mail: urvest@mail.ru
в редакцию журнала «Вестник УрФО. Безопасность в информационной сфере».

Или по почте по адресу: Россия, 454080, г. Челябинск, пр. им. Ленина, д. 76, ЮУрГУ,
Издательский центр

ВЕСТНИК УрФО
Безопасность в информационной сфере № 2(52) / 2024
Подписано в печать 16.07.2024. Дата выхода в свет 21.08.2024.
Формат 70×108 1/16. Печать цифровая. Усл.-печ. л. 7,79. Тираж 50 экз.
Заказ 110/356.
Цена свободная.

Отпечатано в типографии Издательского центра ФГАОУ ВО «ЮУрГУ (НИУ)».
454080, г. Челябинск, пр. им. В. И. Ленина, 76, ЮУрГУ, Издательский центр.

Bulletin of the Ural Federal District
Security in the Sphere of Information No. 2(52) / 2024
Signed to print 16.07.2024. Date of publication of the 21.08.2024.
Format 70×108 1/16. Screen printing. Conventional printed sheet 7,79. Circulation – 50 issues.
Order 110/356.
Open price.

Printed in the printing house of the Publishing Center of FGAOU VO «SUSU (NIU)».
SUSU, Publishing Center, 76, Lenina Str., Chelyabinsk, 454080