ISSN 2225-5435

Вестник УрФО.

БЕЗОПАСНОСТЬ В ИНФОРМАЦИОННОЙ СФЕРЕ

№ 4(54) / 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

№ 4(54) / 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

РАДИОТЕХНИКА, В ТОМ ЧИСЛЕ СИСТЕМЫ И УСТРОЙСТВА ТЕЛЕВИДЕНИЯ

СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ

МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

IN THIS ISSUE

RADIO ENGINEERING, INCLUDING TELEVISION SYSTEMS AND DEVICES

SYSTEM ANALYSIS, MANAGEMENT AND INFORMATION PROCESSING

METHODS AND SYSTEMS OF INFORMATION PROTECTION, INFORMATION SECURITY

УДК 519.254

Вестник УрФО № 4(54) / 2024, с. 5–14

Портнов А. В., Даровских С. Н., Николаев А. Н., Николаева А. Р.

10.14529/secur240401

ОСОБЕННОСТИ ПРИМЕНЕНИЯ АЛГОРИТМА ОПРЕДЕЛЕНИЯ ГРАНИЦ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ ПРИ ИСПОЛЬЗОВАНИИ ГРАДИЕНТНЫХ МЕТОДОВ

В статье обсуждается задача определения границ объектов в автономных системах обработки видеоизображений. Особенностью таких систем являются ограниченные возможности в части вычислительных ресурсов. Предлагаемый подход основан на применении градиентных методов и пространственной фильтрации. Точность детектирования границ объектов определяется выбором наилучшего порогового значения принятия решения относительно их наличия или отсутствия. Существует большое количество градиентных методов детектирования границ объектов, основанных на применении специальных операторов (Робертса, Прюитта, Собеля, Лапласа и др.). В основе большинства алгоритмов градиентных методов лежит пространственная фильтрация, за которой следует процедура бинаризации с применением порогового значения. Однако задача выбора порогового значения в открытой литературе освещена недостаточно. Проведенные исследования показали возможность определения порогов, обеспечивающих минимальную ошибку для различных уровней размытия границы. По полученным экспериментальным данным ставится задача разработки алгоритмов адаптивного определения порогового значения.

Ключевые слова: *градиентные методы, пространственная фильтрация, определение границ, искусственные фасеточные глаза, техническое зрение*

FEATURES OF THE APPLICATION OF THE ALGORITHM FOR DETERMINING THE BOUNDARIES OF OBJECTS IN IMAGES USING GRADIENT METHODS

The article discusses the problem of determining the boundaries of objects in autonomous video image processing systems. A feature of such systems is their limited capabilities in terms of computing resources. The proposed approach is based on the use of gradient methods and spatial filtering. The accuracy of detecting the boundaries of objects is determined by choosing the best threshold value for making a decision about their presence or absence. There are a large number of gradient methods for detecting the boundaries of objects based on the use of special operators (Roberts, Prewitt, Sobel, Laplacian, etc.). Most algorithms of gradient methods are based on spatial filtering, followed by a binarization procedure using a threshold value. However, the problem of choosing a threshold value is not sufficiently covered in the open literature. The conducted studies have shown the possibility of determining thresholds that provide a minimum error for various levels of boundary blurring. Based on the experimental data obtained, the task is to develop algorithms for adaptive threshold value determination.

Keywords: *gradient methods, spatial filtering, boundary detection, artificial faceted eyes, technical vision*

Введение

Задача точного определения координат движущихся объектов является актуальной для многих областей, таких как автономные транспортные средства, робототехника, системы видеонаблюдения, охранные системы и др. Традиционные радиотехнические методы определения координат объектов предполагают использование одно- или многопозиционных активных радиолокационных станций [1]. Наряду с этим широкое распространение получили и пассивные методы определения координат объектов, основанные в том числе на обработке видеоинформации. Такие методы активно используются в автономных робототехнических комплексах с системами машинного зрения. В настоящее время в этих системах широко используются методы обработки, основанные на глубоком обучении нейронных сетей и демонстрирующие высокую точность и гибкость. Однако высокая вычислительная сложность делает невозможным их применение в автономных и компактных робототехнических комплексах. Это обуславливает необходимость разработки эффективных алгоритмов обработки видеоданных, не требующих больших вычислительных затрат.

Для сокращения требуемого объема вычислений предложен способ обработки изображений, основанный на принципах фасеточного зрения, заимствованного у живой природы [2]. В качестве алгоритма выделения объекта, его границ предлагается использование градиентных методов и пространственной фильтрации изображений [3]. Их простота реализации и низкая вычислительная сложность позволяют эффективно обра-

батывать видеопоток в режиме реального времени на маломощных платформах. Однако, точность работы алгоритмов градиентных методов сильно зависит от ряда факторов, таких как уровень шума на изображении, освещенность сцены и т.д. Для точного определения реальных границ и минимизации ложных, необходимо подбирать пороговое значение принятия решения относительно наличия или отсутствия границ. Некорректное определение порога приводит к неполному или, наоборот, к ложному выделению границ.

Таким образом, возникает проблема выбора оптимального порогового значения, обеспечивающего баланс между полнотой выделения границ и их пропуском. В данной статье рассматриваются особенности градиентных методов и проводится анализ влияния степени размытия границ на точность определения контура объекта. Цель исследования заключается в определении возможности выбора наилучшего порогового значения, обеспечивающего минимальную ошибку определения контура для различных уровней размытия (яркости) границ. В основе выбора порогового значения лежат особенности использования градиентных методов и алгоритмов пространственной фильтрации.

Градиентные методы и пространственная фильтрация

Детектирование границ с помощью градиентных методов – это процесс определения изменений интенсивности пикселей. Они используются для обнаружения и распознавания объектов, сегментации изображений и других задач. В данной статье рассматривается метод выделения границ, основанный на применении оператора Собеля, который является компромиссным методом выделения границ с точки зрения соотношения качества и вычислительных затрат [4].

Градиентные методы выделения границ основаны на пространственной фильтрации. Так как мы оперируем изображением, пространственной областью является плоскость изображения, состоящая из пикселей. В общем виде процесс пространственной фильтрации представляет собой применение того или иного оператора к изображению (рис. 1).

На рис. 1 представлено исходное изображение, на котором выбрана рассматриваемая точка с координатами (x, y) и ее окрестность, представленная в виде прямоугольной области вокруг нее. После применения какого-либо оператора к выбранной окрестности точки, полученный результат записывается в новую матрицу, после чего рассматривается соседняя точка, например, с координатами $(x+1, y)$. Процесс повторяется до тех пор, пока не будут рассмотрены все точки изображения.

Результатом наложения оператора (ядра, маски фильтра) на окрестность выбранной точки изображения (рис. 2) является сумма произведений соответствующих друг другу пикселей и коэффициентов фильтра в ядре.

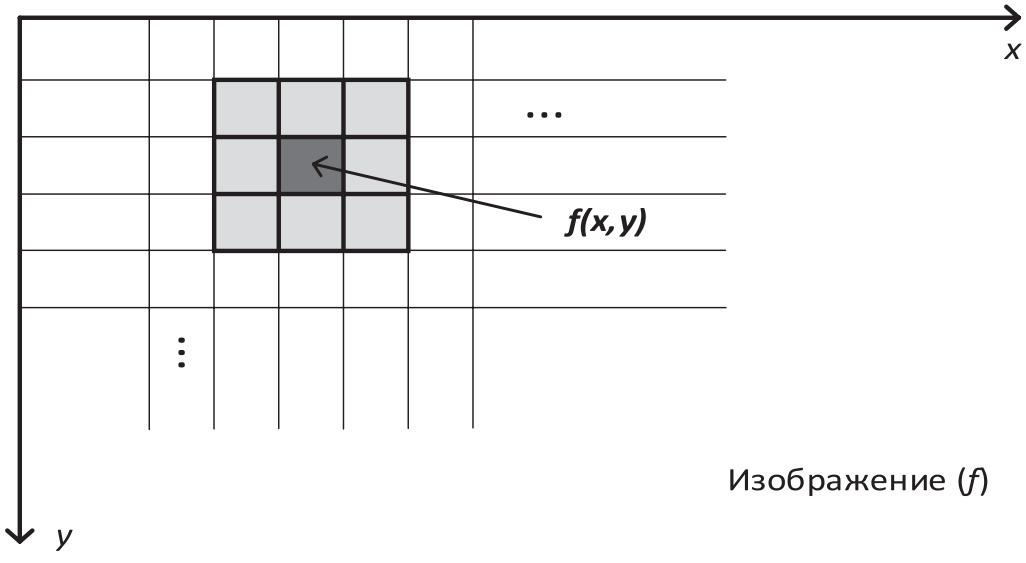


Рис. 1. Пояснение метода пространственной фильтрации

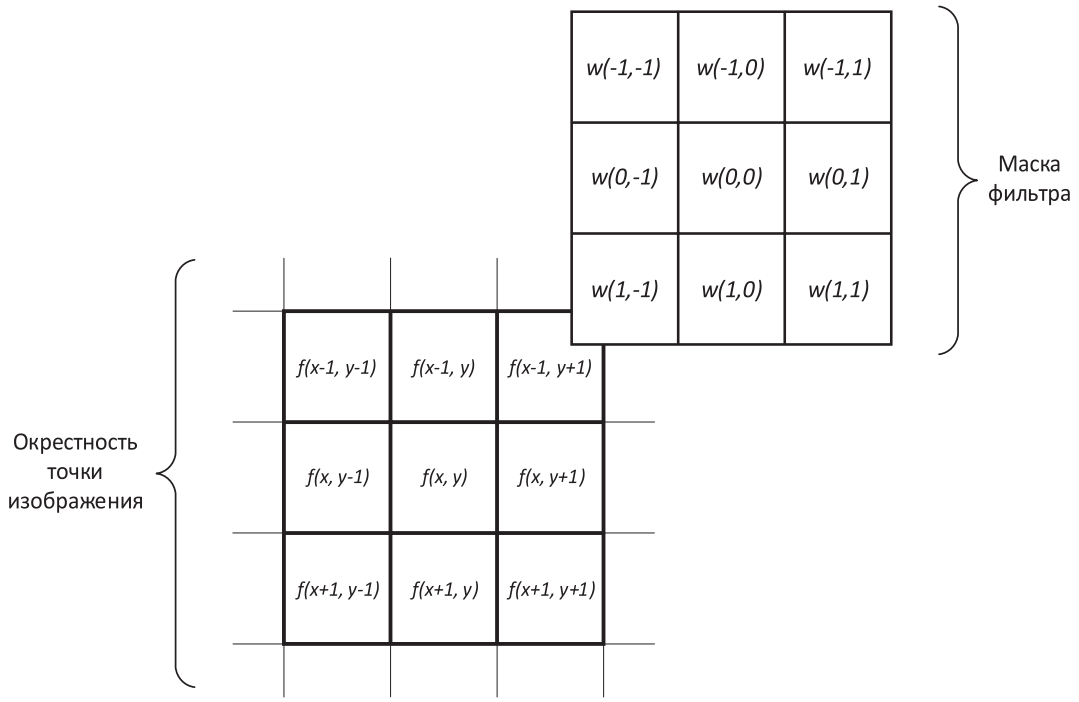


Рис. 2. Наложение ядра фильтра на окрестность выбранной точки

Математически процесс фильтрации можно представить с помощью выражения [3]:

$$g(x,y) = \sum_{s=\frac{m-1}{2}}^{\frac{m-1}{2}} \sum_{t=\frac{n-1}{2}}^{\frac{n-1}{2}} w(s,t) f(x+s, y+t),$$

где m и n – размеры маски фильтра, $w(s,t)$ – коэффициенты маски, f(x,y) – значения пикселей в окрестности выбранной точки.

Для оператора размером m=n=3 результат свертки рассматриваемой области изображения с ядром фильтра можно представить в виде:

$$g(x,y) = w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + \cdots +$$
$$+w(0,0)f(x,y) + \cdots + w(1,0)f(x+1,y) + w(1,1)f(x+1,y+1).$$

Маски фильтров выделения границ

В общем случае размеры масок могут быть произвольными, но чаще всего используются маски с нечетными размерами (в основном 3x3), так как они являются центрированными. При их использовании рассматриваются все пиксели вокруг выбранной точки. Коэффициенты ядер также могут быть различными, они напрямую влияют на результат фильтрации [3].

Градиентные методы выделения границ основаны на вычислении производной интенсивности изображения по горизонтали и вертикали. Градиент указывает направление и величину наибольшего изменения функции (в данном случае, интенсивности изображения) в рассматриваемой точке. Так как изображение является дискретным [5], то частные производные для любой выбранной точки можно получить с помощью выражений [3, 6]:

$$g_x = \frac{\partial f(x,y)}{\partial x} = f(x+1,y) - f(x,y);$$
$$g_y = \frac{\partial f(x,y)}{\partial y} = f(x,y+1) - f(x,y).$$

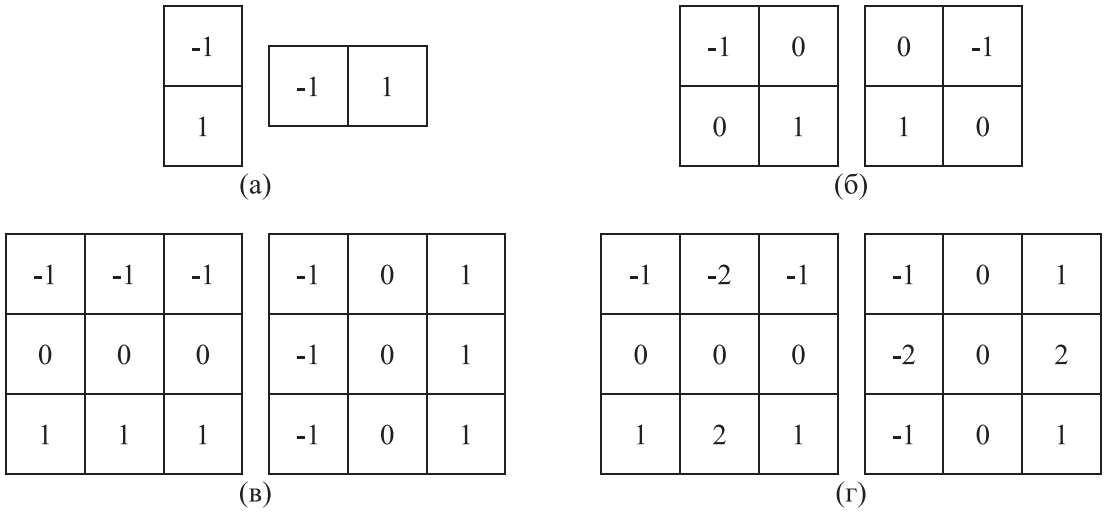


Рис. 3. Распространенные маски фильтра: (а) одномерные маски; (б) маски оператора Робертса; (в) маски оператора Прюитта; (г) маски оператора Собеля

Они позволяют определять только изменение яркости в горизонтальном и вертикальном направлениях. В данном случае маски фильтра будут являться одномерными (рис. 3, а).

Для определения изменения яркости в двух диагоналях одномерные маски модифицируются до двумерных (рис. 3, б). Данные ядра известны как маски оператора Робертса. Как видно, маски размером 2x2 не являются центрированными и учитывают не все пиксели вокруг рассматриваемой точки. Для этого необходимы маски с нечетными размерами. Самыми простыми являются маски оператора Прюитта [3, 4], они имеют размер 3x3 (рис. 3, в). Такие ядра являются симметричными относительно центрального значения, что позволяет оперировать всеми значениями интенсивности пикселей вокруг центральной точки.

Для уменьшения эффекта сглаживания [3], а следовательно, усиления градиента между пикселями, средние боковые коэффициенты ядер могут быть удвоены, в результате чего получаются маски, известные как оператор Собеля (рис. 3, г) [7]. Удвоенные коэффициенты масок фильтра способствуют большей чувствительности к изменениям интенсивности вдоль направления вычисления производной.

Особенности выбора порогового значения фильтра

В качестве исходного изображения был использован простой переход с темного цвета на светлый (рис. 4, а). Данное изображение является идеальным (не имеет шума) и содержит одну границу между двумя областями (темной и светлой), которая четко определена и легко верифицируема. Это позволяет

(а) (б)

Рис. 4. Тестовое изображение: (а) исходное изображение; (б) истинная граница

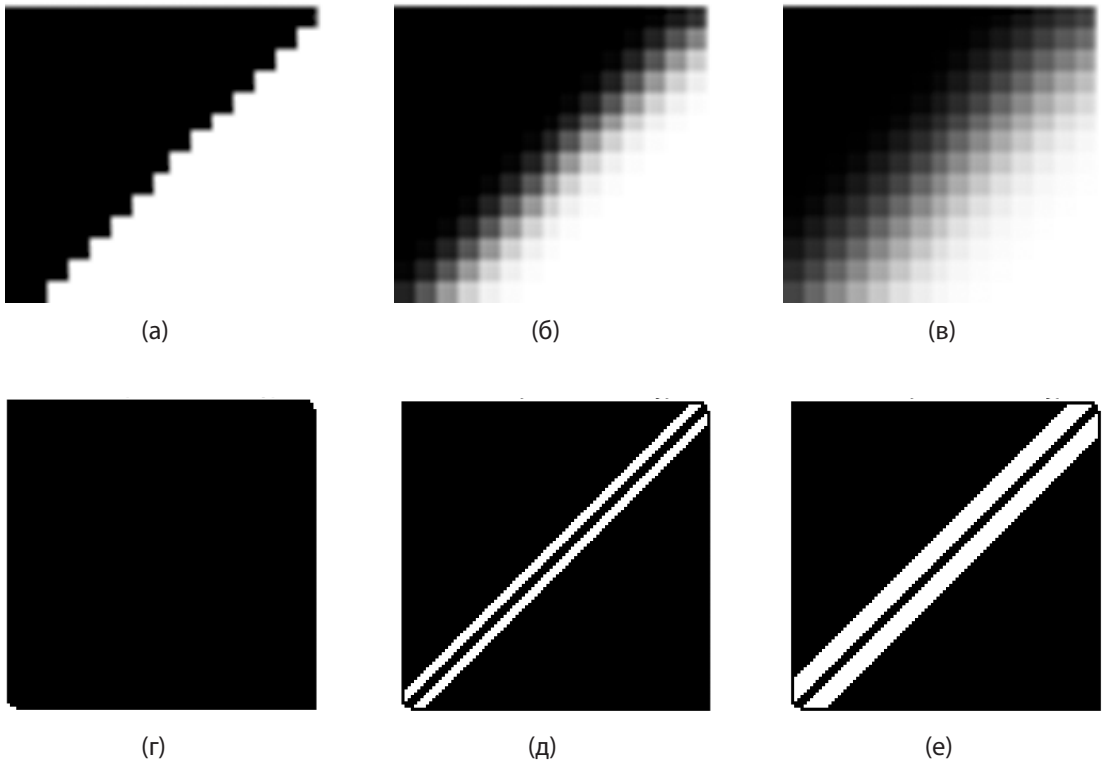


Рис. 5. Определение границы исследуемого изображения: (а) четкая граница;
(б), (в) размытая с разной степенью граница;
(г), (д), (е) ошибка определения (белые пиксели) границы для (а), (б), (в) соответственно

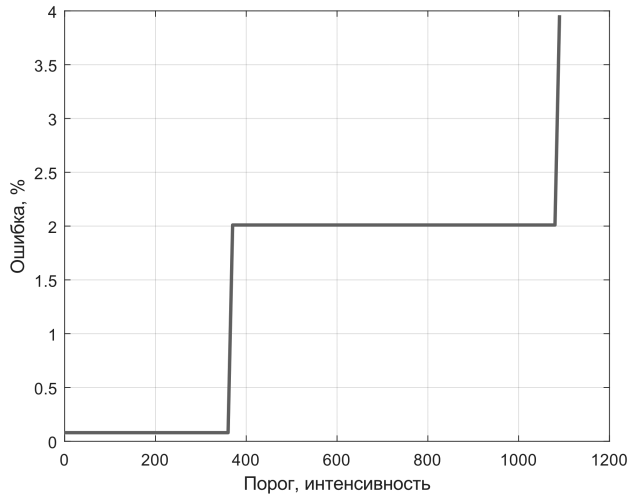
определить истинную границу (рис. 4, б) для дальнейшего сравнения результатов детектирования контура. В данном случае для выделения границы используется оператор Собеля.

Для исследования особенностей выбора порогового значения рассматриваются три случая (рис. 5, а-в) с различной степенью размытия границы (для наглядности представлена часть изображения в увеличенном масштабе).

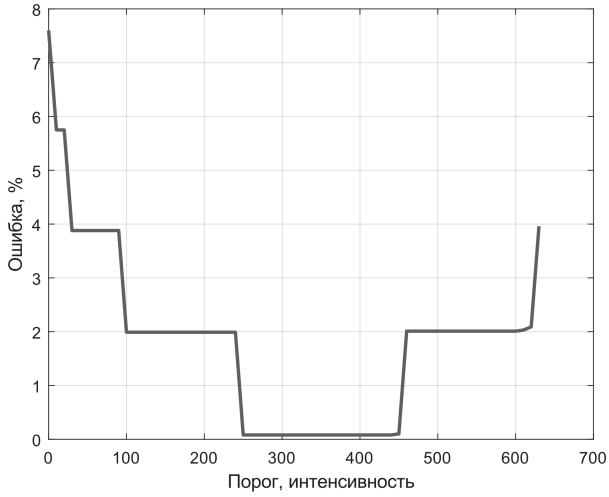
Результат работы фильтра (бинарное изображение после пороговой обработки) сравнивается поэлементно с эталонным бинарным изображением с помощью операции исключающего «ИЛИ» (XOR). Операция XOR возвращает единицу, если пиксели в двух изображениях имеют разные значения (один белый, другой черный), и ноль, если эти значения одинаковые. Таким образом, результирующее изображение после операции XOR покажет все позиции, в которых граница не определилась, или определилась неверно. Белые пиксели на изображении (рис. 5, г-е) представляют собой ошибку определения границы (чем шире область, обозначенная белым цветом, тем больше ошибка).

Для количественной оценки точности определения границы изображения подсчитывается общее количество белых пикселей, полученных после операции XOR. Это количество делится на общее число пикселей в изображении и выражается в процентах. Процент ошибки отражает долю пикселей, в которых используемый оператор дал неверный результат. Для анализа влияния порогового значения фильтра на точность детектирования границы объекта строится график зависимости процента ошибки от значения порога для разной степени размытия изображения (рис. 6).

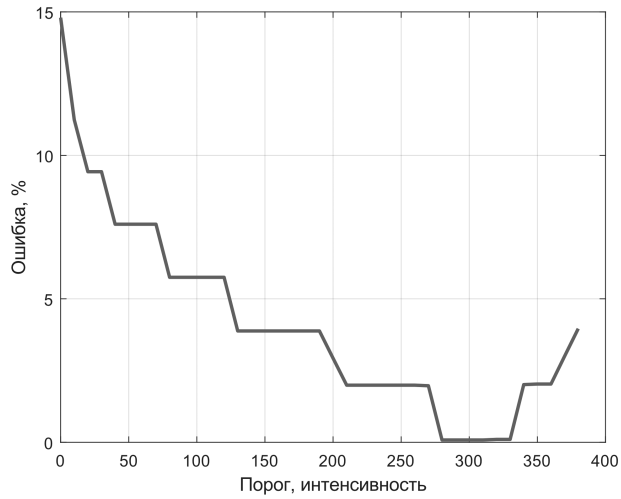
Пороговое значение принятия решения относительно наличия или отсутствия границы определяется требованиями к ширине границы выходного бинарного изображения (чем больше порог, тем уже граница). Оно определяет чувствительность алгоритма к изменениям яркости на изображении – низкий порог приводит к детектированию большего количества границ (включая ложные), а высокий порог в ряде случаев может приве-



(а)

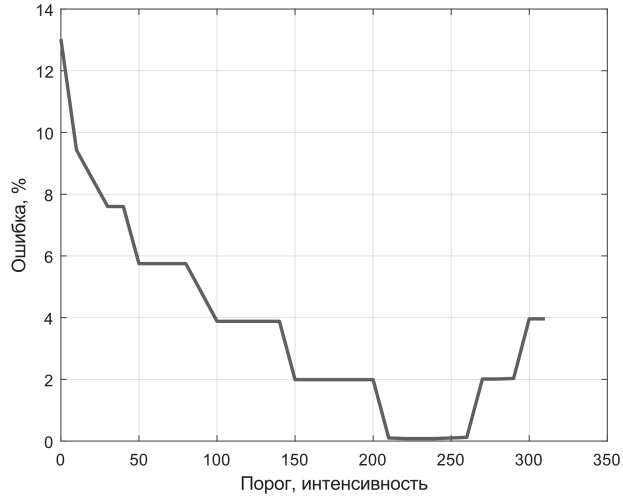


(б)

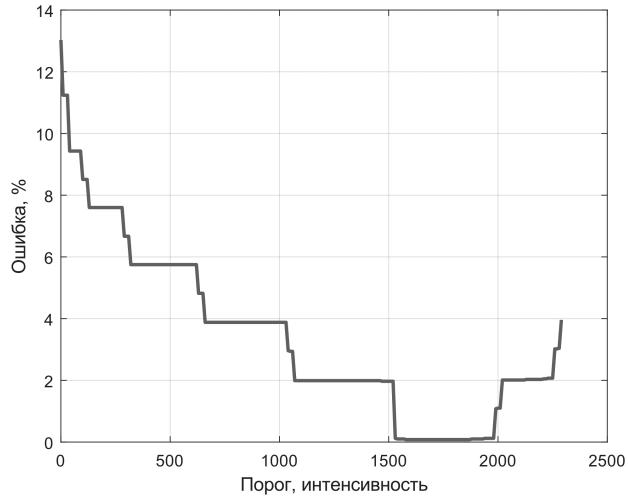


(в)

Рис. 6. Зависимость ошибки определения границ от порогового значения: (а) четкая граница; (б) слабое размытие границы; (в) сильное размытие границы



(а)



(б)

Рис. 7. Зависимость ошибки определения границ от порогового значения:
(а) оператор Прюитта; (б) оператор Собеля с центральными элементами, увеличенными в 10 раз

сти к пропуску реальных границ. Анализ полученных данных (рис. 6) показывает реальную зависимость точности детектирования границы объекта от яркости изображения. Оно позволяет сделать вывод о том, что для различных ситуаций возможно определить пороговое значение, при котором процент ошибки определения границ минимален. Он будет зависеть от характеристик исходного изображения и может изменяться для различных ситуаций (освещение, зашумленность и т.д.).

На рис. 7 представлены графические зависимости ошибки определения границы объекта от порогового значения при использовании фильтра с другими ядрами. Рассматривается случай с сильно размытой границей.

Анализ полученных зависимостей (рис. 6, в и рис. 7, а-б) позволяет сделать вывод о том, что в количественном плане при данном алгоритме расчета, график ошибки имеет схожий вид. Различия заключаются в уменьшении или увеличении пороговой величины для сохранения минимальной ошибки. Таким образом, с увеличением веса центрального элемента в ядрах фильтра увеличивается «окно», в котором должно находиться пороговое значение.

Заключение

Использование простого тестового изображения с четким и плавным диагональным переходом позволило исключить влияние шума и других явлений, сосредоточившись на оценке работы фильтра выделения границ. Полученные результаты, представленные в виде графических зависимостей ошибки выделенных границ от порогового значения фильтра, наглядно демонстрируют его влияние на точность определения контура объекта. Анализ полученных результатов позволяет определить пороговое значение для конкретного типа изображения. Выбор этого значения является компромиссным. Он должен обеспечить качественное определение реальных границ объекта и подавлять ложные контуры, возникающие при наличии шума и других искажений изображений.

Тип масок фильтра также влияет на выбор порога. Увеличение веса центральных элементов в ядрах фильтра расширяет область оптимальных значений порога. Результаты данного исследования могут быть использованы при разработке алгоритмов первичной обработки видеоинформации в системах технического зрения на основе искусственных фасеточных глаз [2, 8, 9]. Вместе с тем результаты исследования показывают, что требуется проведение дополнительных исследований для создания алгоритмов, реализующих адаптивное определение порогового значения для оценки качества того или иного типа изображения.

Литература

1. Фарина А. Цифровая обработка радиолокационной информации. Сопровождение целей. / А. Фарина, Ф. Студер. – М.: Изд-во «Радио и связь», 1993. – 315 с.

2. Портнов А.В., Николаев А.Н., Николаева А.Р. Устройство и алгоритмы первичной обработки видеоинформации для системы машинного зрения на основе искусственных фасеточных глаз // Вестник УрФО. Безопасность в информационной сфере, 2023. – № 4 (50). – С. 5-12. DOI: 10.14529/secur230401.

3. Гонсалес Р. Цифровая обработка изображений: Издание 3-е, исправленное и дополненное / Р. Гонсалес, Р. Вудс // Москва: Техносфера, 2012. – 1104 с.

4. Пахомова О.А. Сравнительный анализ градиентных методов выделения контура объекта на изображении / О.А. Пахомова, О.Я. Кравец // Вестник ВГТУ, 2018. – № 1(14). – С. 37-43.

5. Прэтт У. Цифровая обработка изображений. Пер. с англ. / У. Прэтт // М.: Мир, 1982. – 480 с.

6. Методы компьютерной обработки изображений / М.В. Гашников, Н.И. Глумов, Н.Ю. Ильясова [и др.]. – Под ред. В.А. Сойфера. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 784 с. – ISBN 5-9221-0270-2.

7. Sobel I. A 3x3 Isotropic Gradient Operator for Image Processing / I. Sobel, G. Feldman // Pattern Classification and Scene Analysis, 1968. – P. 271-272.

8. Соломатин В. Фасеточное зрение: перспективы в оптико-электронных системах // Фотоника, 2009. – № 1. – С. 22–26.

9. Федянина Р.С., Соколинский Л.Б. Двумерная модель фасеточного зрения // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика, 2020. – Т. 9, № 4. – С. 48-66. DOI: 10.14529/cmse200404.

References

1. Farina A. Tsifrovaya obrabotka radiolokatsionnoy informatsii. Soprovozhdeniye tseley. / A. Farina, F. Studer. – M.: Izd-vo «Radio i svyaz'», 1993. – 315 s.

2. Portnov A.V., Nikolayev A.N., Nikolayeva A.R. Ustroystvo i algoritmy pervichnoy obrabotki videoinformatsii dlya sistemy mashinnogo zreniya na osnove iskusstvennykh fasetochnykh glaz // Vestnik UrFO. Bezopasnost' v informatsionnoy sfere, 2023. – № 4 (50). – S. 5-12. DOI: 10.14529/secur230401.

3. Gonsales R. Tsifrovaya obrabotka izobrazheniy: Izdaniye 3-ye, ispravlennoye i dopolnennoye / R. Gonsales, R. Vuds // Moskva: Tekhnosfera, 2012. – 1104 s.

4. Pakhomova O.A. Sravnitel'nyy analiz gradiyentnykh metodov vydeleniya kontura ob"yekta na izobrazhenii / O.A. Pakhomova, O.YA. Kravets // Vestnik VGTU, 2018. – № 1(14). – S. 37-43.

5. Prett U. Tsifrovaya obrabotka izobrazheniy. Per. s angl. / U. Prett // M.: Mir, 1982. – 480 s.

6. Metody komp'yuternoy obrabotki izobrazheniy / M.V. Gashnikov, N.I. Glumov, N.YU. Il'yasova [i dr.]. – Pod red. V.A. Soyfera. – 2-ye izd., ispr. – M.: FIZMATLIT, 2003. – 784 s. – ISBN 5-9221-0270-2.

7. Sobel I. A 3x3 Isotropic Gradient Operator for Image Processing / I. Sobel, G. Feldman // Pattern Classification and Scene Analysis, 1968. – p. 271-272.

8. Solomatin V. Fasetochnoye zreniye: perspektivy v optiko-elektronnykh sistemakh // Fotonika, 2009. – № 1. – S. 22–26.

9. Fedyanina R.S., Sokolinskiy L.B. Dvumernaya model' fasetochnogo zreniya // Vestnik YUUrGU. Seriya: Vychislitel'naya matematika i informatika, 2020. – T. 9, № 4. – S. 48-66. DOI: 10.14529/cmse200404.

Портнов Андрей Владимирович, аспирант ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: portnovav@susu.ru

Даровских Станислав Никифорович, доктор технических наук, доцент, профессор ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: darovskikhsn@susu.ru

Николаев Андрей Николаевич, доцент ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: nikolaevan@susu.ru

Николаева Алиса Робертовна, инженер ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: nikolaevaar@susu.ru

Portnov Andrey Vladimirovich, post-graduate student of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: portnovav@susu.ru

Darovskikh Stanislav Nikiforovich, doctor of technical sciences, docent, professor of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: darovskikhsn@susu.ru

Nikolaev Andrey Nikolaevich, docent of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: nikolaevan@susu.ru

Nikolaeva Alisa Robertovna, engineer of the Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: nikolaevaar@susu.ru

УДК 004.832.2 Вестник УрФО № 4(54) / 2024, с. 15–21

Кротова Е. Л., Субботина Ю. В., Ермаков Д. Г., Тишин К. Л.

10.14529/secur240402

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ БЛОКЧЕЙН ПРИ РАЗРАБОТКЕ СИСТЕМЫ ЭЛЕКТРОННОГО ГОЛОСОВАНИЯ

Проведен сравнительный анализ самых востребованных блокчейн проектов на применимость для решения задачи разработки протокола электронного голосования. Выделены те принципы технологии blockchain, которые позволяют улучшить существующие решения в области электронного голосования Разработаны критерии оценки надежности разрабатываемого протокола. Показано, что технология blockchain и принципы, на которых она базируется, позволяют использовать ее положительные качества для создания электронной системы голосования, в которой предусмотрено устранение недостатков существующих систем электронного голосования, сформулированных ранее.

Ключевые слова: *технология блокчейн, P2P протоколы, электронное голосование, КОИБ*

Krotova E. L., Subbotina Yu. V., Ermakov D. G., Tishin K. L.

USING BLOCKCHAIN TECHNOLOGY TO DEVELOP AN ELECTRONIC VOTING SYSTEM

A comparative analysis of the most popular blockchain projects was carried out to determine their applicability for solving the problem of developing an electronic voting protocol. The principles of blockchain technology are highlighted that make it possible to improve existing solutions in the field of electronic voting. Criteria for assessing the reliability of the protocol being developed have been developed. It is shown that blockchain technology and the principles on which it is based make it possible to use its positive qualities to create an electronic voting system, which provides for the elimination of the shortcomings of existing electronic voting systems formulated earlier.

Keywords: *blockchain technology, P2P protocols, e-voting, PECs*

Понятие «электронное голосование» можно определить как совокупность различных способов выражения воли избирателя, объединенных обязательным условием: подсчет голосов осуществляется с помощью специальных программно-технических устройств без вмешательства человека. В постановлении ЦИК России от 27 августа 2014 года № 248/1529–6 «О Порядке электронного голосования с использованием комплексов для электронного голосования на выборах, проводимых в Российской Федерации» дается следующее определение: «Электронное голосование – голосование без использования бюллетеня на бумажном носителе, с использованием комплекса средств автоматизации ГАС «Выборы». Таким образом, в России электронное голосование не включает использование оптических машин для сканирования бумажных бюллетеней.

В день голосования избиратель имеет право выбрать способ участия в выборах: дистанционно (электронно) или традиционно (на избирательном участке).

Анализируя мировой опыт электронного голосования, можно отметить, что в большинстве стран электронное голосование охватывает любую форму выражения воли граждан, где для обработки голосов используются программно-технические средства. Формы электронного голосования могут быть различными, но процесс передачи информации – протокол – един.

Выборы остаются одной из проблемных тем в общественном восприятии. Итоги, процесс голосования, свобода выбора, анонимность – все это вызывает сомнения. В связи с этим вырос интерес к новым технологиям тайного голосования, особенно к дистанционному электронному голосованию (ДЭГ).

Впервые ДЭГ было применено в единый день голосования 8 сентября 2019 года в Москве на выборах депутатов городской думы. Оно проводилось как эксперимент в трех округах столицы. В 2020 году ДЭГ использовалось в общероссийском голосовании по поправкам в Конституцию РФ в Москве и Нижегородской области. В 2021 году ДЭГ применялось на выборах в Госдуму РФ в семи регионах, в 2022 году – в восьми регионах, а в 2023 году – уже в 25 регионах.

ДЭГ имеет несколько преимуществ по сравнению с традиционной системой голосования:

Удобство для избирателей: возможность голосовать в любом месте, что особенно важно для маломобильных граждан или тех, кто находится не по месту прописки.

Увеличение явки: избиратели могут голосовать, не выходя из дома.

Исключение человеческого фактора: автоматический подсчет голосов обеспечивает большую точность и скорость.

Снижение затрат: отсутствие необходимости в привлечении большого числа людей для организации процесса и аренде специальных площадок делает выборы дешевле и проще в подготовке.

Основным препятствием для широкого внедрения электронного голосования остается стереотип о ненадежности результатов информационных систем. Люди часто не доверяют результатам, полученным через Интернет. В данной работе рассматриваются вопросы обеспечения информационной безопасности электронных выборов с использованием протоколов с нулевым разглашением.

Объектом исследования в данной теме являются протоколы с нулевым разглашением и протоколы на основе технологии блокчейн.

Предметом исследования является определение области использования электронного голосования, построенного на основе протоколов с нулевым разглашением и на основе технологии блокчейн.

Научно-техническая проблема определена тем, что в условиях активного развития информационных технологий и повышенного интереса граждан к анонимности, важной проблемой для отрасли становится приобретение статуса конфиденциальной информации при реализации конституционного права свободного выбора. Одним из наиболее эффективных методов решения проблемы является усовершенствование системы с использованием протоколов с нулевым разглашением и на основе технологии блокчейн.

Исследуя возможность использования технологии блокчейн для электронного голосования было выявлено, что основной характеристикой программного обеспечения является уровень доверия в системе.

Доверие к системе основывается на вычислительной работе хэш-функции, которая преобразовывает данные в код определенной длины, т.е. шифрует записи. С помощью полученного кода определенной длины третья сторона понимает, что данные,

которые были переданы достоверны, при этом даже не зная самих данных. Аргументом такой функции может быть сообщение или файл, а значение - цепочка битов определенной длины, например, 256 бит. Выходное значение функции называется «хэшем» или «дайджестом» сообщения. Значение всегда одно и то же для определенного аргумента. Если изменится хотя бы один символ в аргументе, то полученный хэш будет совершенно другим. Так как функция является криптографической, то ее расчет в обратном направлении требует невыполнимых вычислений [1-7]. Мы исключаем возможность принятия поддельных операций и реестров с несоответствиями, так как они требуют практические неосуществимых по времени вычислений.

Допустим, что в распределенном реестре есть 4 участника: Мы, Алиса, Боб, и Чарли. Все транслируют друг другу сообщения об операциях и нам нужно каким-то образом, прийти к согласию насчет правильного реестра, который каждый из участников транслирует. Сначала реестр делится на блоки, каждый из которых состоит из списка операций и доказательства выполнения, то есть числа, вместе с которыми хэш блока начинается с какого-то количества нулей. Блок действителен, только если в нем есть доказательства выполнения работы, аналогично операции, которая подтверждается только если ее подписывает отправитель. Мы доверяет тому реестру, над которым было проведено больше всего вычислений. Если они одинаковые по длине, то нужно дождаться нового блока, который изменит это равенство.

Также для того, чтобы мы, Алиса, Боб и Чарли достигли согласия о текущем состоянии данных во всех блоках, в блокчейн встраивается специальный механизм под названием алгоритм консенсуса. Соблюдение правил этого алгоритма дает гарантию того, что все транзакции достоверны. То есть алгоритм консенсуса говорит нам о том, что все участники (мы, Алиса, Боб, Чарли), подключенные к блокчейну были согласны с добавлением нового блока в цепочку. Наиболее востребованными среди лучших блокчейн-проектов являются: Proof-of-Work, Proof-of-Stake, PoA, DPoS.

Говоря об идеальной системе голосования стоит заметить, что достичь ее невозможно, но рассмотрим, каким требованиям должна удовлетворять наша система [5].

1. Все голоса должны быть учтены и учтены корректно.

2. Верифицируемость избирателей. Возможность голосования должна быть предоставлена всем лицам, обладающим избирательным правом.

3. Один избиратель – один голос, не должно допускаться двойное голосование.

4. Голосование должно быть анонимным.

5. Голосование должно исключать возможность какого-либо контроля за волеизъявлением гражданина или принуждения к нему.

6. Голосование не должно создавать предпосылок для манипуляций на основе политических взглядов.

7. Должна обеспечиваться неизменность поданного голоса. К тому же данные о волеизъявлении избирателей не могут быть изменены или удалены.

8. Процесс голосования должен быть прозрачен, поскольку цель выборов – не просто выбрать победителя, но и убедить проигравших в том, что они проиграли.

9. Конфиденциальность голосов. Должна отсутствовать возможность подсчета промежуточных результатов голосования до его завершения. Конфиденциальность достигается за счет шифрования бюллетеней и невозможности расшифрования до окончания голосования.

10. Должна обеспечиваться прозрачность исполнения и неизменность программного кода, реализуемого в виде смарт-контрактов.

11. Должна обеспечиваться защита и неизменность данных, используемых в процессе голосования: списка избирателей, ключах, используемых для шифрования бюллетеней на различных этапах криптографического протокола, и так далее.

12. Децентрализованное хранение данных, при этом каждый участник должен обладать абсолютно идентичной со всеми копией, подтвержденной свойствами консенсуса в сети. [6]

13. Должна обеспечиваться возможность просматривать транзакции и отслеживать ход голосования, полностью отражающегося в цепочках блоков, от его начала до записи рассчитанных итогов.

14. Проверяемость. Наблюдатель может проверить, что подсчет голосов осуществлялся корректным образом.

Ключевой критерий оценки успешности избирательной системы – это высокий уровень доверия со стороны граждан. Именно этот критерий позволяет избежать волнения среди населения в пост-выборный период и дает возможность стране перейти в будущее без серьезных потрясений. Таким образом, чтобы сделать успешную избирательную систему, нужно добиться максимальной прозрачности как в бизнес-процессах, так и в технических аспектах. Однако при абсолютной анонимности, прозрачность теряется. [7]

Принципы технологии блокчейн

В этом разделе подведем итог и обозначим те принципы технологии blockchain, которые позволяют улучшить существующие решения в области электронного голосования.

• Копии с информацией о транзакциях распространяются среди большого количества участников системы для того, чтобы избежать критических ошибок в случае, если что-то произойдет на одном устройстве.

Этот принцип позволяет соблюдать критерий достаточной отказоустойчивости системы голосования.

• Невозможность внесения изменений или уничтожения записей в блокчейн без консенсуса.

• Использование хэширования для защиты данных и проверки подлинности отправителя и получателя.

Эти принципы позволяют соблюдать критерий надежности системы к взлому или искажению информации третьими лицами.

• Открытый публичный бухгалтерский реестр, который позволяет получить доступ к деталям транзакции с момента создания блокчейна, никак не раскрывая личностей людей, которые участвовали в этих транзакциях.

Этот принцип позволяет соблюдать критерий достаточной прозрачности процесса голосования, который мы выделили раннее для лучшего функционирования системы электронного голосования.

Требования к системе голосования с применением технологии блокчейн

Требования, которые должны быть реализованы для системы голосования на блокчейн:

1. Возможность создания опросов.

После создания пользователем опроса, ему будет присуждаться статус администратора опроса. После чего он будет иметь возможность разграничивать доступ для лиц, которые будут иметь разрешение на участие в созданном опросе (голосовании).

2. Возможность создания списков объектов голосования к опросу.

3. Возможность зарегистрировать участников созданного опроса.

4. Возможность для участников опроса голосовать.

Каждый пользователь при успешном прохождении регистрации должен получать необходимое для подачи голоса число токенов.

5. Обеспеченная прозрачность процесса голосования.

Каждый пользователь должен иметь возможность посмотреть результаты голосования и цепочки блоков в реальном времени.

6. Обеспечение отказоустойчивости системы.

Система продолжит работу, если одно из устройств с базой данных голосования выйдет из строя, так как копия базы данных хранится на устройствах всех участников сети, где запущено децентрализованное приложение.

7. Отсутствие возможности вносить любые несанкционированные изменения, влияющие на подсчет голосов

У злоумышленника должна отсутствовать возможность влияния на ход голосования и его результаты из-за устойчивости системы к взлому.

В ходе аналитической работы был выявлен ряд целей, к которым стремиться любая система электронного голосования:

- повышение скорости подсчета голосов;
- сокращение роли человека в подсчете голосов;
- повышение доверия к результатам голосования.

С опорой на успешный мировой опыт автоматизации голосования и на характеристики/критерии, следование которым свидетельствовало бы об улучшении текущих решений, предложено следующее решение сформулированной проблемы: использование технологии blockchain при разработке системы электронного голосования. Так как технология blockchain и принципы, на которых она базируется, позволяют использовать ее положительные качества для создания

электронной системы голосования, в которой предусмотрено устранение недостатков существующих систем электронного голосования, сформулированных ранее. Данное решение соответствует всем критериям более эффективной системы электронного голосования, которые были выявлены в ходе анализа, а именно:

• Критерий достаточной прозрачности процесса голосования. Подсчет голосов должен осуществляться корректным образом без возможности его фальсификации.

• Критерий достаточной отказоустойчивости системы голосования. Недостаточность отказоустойчивости системы, может привести к тому, что если машина, занимающаяся обработкой голосов, выйдет из строя, то это приведет к нарушению всего процесса голосования.

• Критерий надежности системы голосования. Принцип построения архитектуры системы голосования должен обеспечивать отсутствие единой «точки отказа» этой системы.

С опорой на успешный мировой опыт автоматизации голосования [4, 5, 7, 10, 12] и на характеристики/критерии, следование которым свидетельствовало бы об улучшении текущих решений, предложено следующее решение сформулированной проблемы: использование технологии blockchain при разработке системы электронного голосования. Так как технология blockchain и принципы, на которых она базируется, позволяют использовать ее положительные качества для создания электронной системы голосования, в которой предусмотрено устранение недостатков существующих систем электронного голосования, сформулированных ранее. Данное решение соответствует всем критериям более эффективной системы электронного голосования, которые были выявлены в ходе изучения теоретических основ технологии блокчейн и ее анализа, а именно:

1. Критерий достаточной прозрачности процесса голосования. Подсчет голосов должен осуществляться корректным образом без возможности его фальсификации.

2. Критерий достаточной отказоустойчивости системы голосования. Недостаточность отказоустойчивости системы, может привести к тому, что если машина, занимающаяся обработкой голосов, выйдет из строя, то это приведет к нарушению всего процесса голосования.

3. Критерий надежности системы голосования. Принцип построения архитектуры системы голосования должен обеспечивать отсутствие единой «точки отказа» этой системы.

Литература

1. Прасти Н. Блокчейн. Разработка приложений, // Н. Прасти, В.С. Яценков. – СПб.: БХВ–Петербург, 2018. – 256 с.

2. Равал С. Децентрализованные приложения. Технология Blockchain в действии, // С. Равал. – СПб.: Питер, – 2017. – 192 с.

3. Тапскотт Д., Тапскотт А. Технология блокчейн – то, что движет финансовой революцией сегодня, // Д. Тапскотт, А. Тапскотт. – М.: Эксмо, 2017. – 448 с.

4. Насколько надежно электронное голосование [Электронный ресурс]. – Режим доступа: https://www.svoboda.org/a/269300.html, свободный.

5. Норвегия официально отказалась от электронного голосования на выборах: оно контрпродуктивно [Электронный ресурс]. – Режим доступа: http://www.mk.ru/politics/world/2014/06/30/norvegiya–otkazalasv–politike–ot–elektronnogo–golosovaniya.html, свободный.

6. Block The Vote: Could Blockchain Technology Cybersecure Elections? [Электронный ресурс]. – Режим доступа: http://www.forbes.com/sites/realspin/2016/08/30/block–the–votecouldblockchain–technology–cybersecure–elections, свободный.

7. California: The Top to Bottom Review [Электронный ресурс]. – Режим доступа: http://www.votetrustusa.org/index.php?option=com_content&task=view&id =2554&Itemid=113, свободный.

8. IGS Votomatic Prototype Goes to the Smithsonian [Электронный ресурс]. – Режим доступа: https://web.archive.org/web/20070713201451/http://www.igs.berkeley.edu/publications/par/winter2001/votomatic.htm, свободный.

9. Kiwi. Bitcoin testnet sandbox. [Электронный ресурс]. – Режим доступа: https://testnet.manu.backend.hamburg/faucet, свободный.

10. NSW election result could be challenged over iVote security flaw [Электронный ресурс]. – Режим доступа: https://www.theguardian.com/australia–news/2015/mar/23/nsw–electionresult–could–be–challenged–over–ivote–security–flaw, свободный.

11. Peer–to–peer [Электронный ресурс]. – Режим доступа: https://bitcoin.org/bitcoin.pdf, свободный.

12. Russian Hackers Acted to Aid Trump in Election, U.S. Says [Электронный ресурс]. – Режим доступа: https://www.nytimes.com/2016/12/09/us/obama–russia–election–hack.html, свободный.

13. Slim. Middeware–slim. [Электронный ресурс]. – Режим доступа: https://www.slimframework.com/docs/v3/concepts/middleware.html, свободный.

14. State bans electronic balloting in 4 counties / Touch–screen firm accused of 'reprehensible,' illegal conduct [Электронный ресурс]. – Режим доступа: https://www.sfgate.com/politics/article/State–bans–electronic–balloting–in–4– counties–2784975.php, свободный.

15. Top 100 Cryptocurrencies by Market Capitalization [Электронный ресурс]. – Режим доступа: https://coinmarketcap.com/, свободный.

16. Voting Machine Company Submits to Inquiry [Электронный ресурс]. – Режим доступа: https://www.nytimes.com/2006/10/31/us/politics/31vote.html?_r=1&oref=slogin, свободный.

17. Why machines are bad at counting votes [Электронный ресурс]. – Режим доступа: https://www.theguardian.com/technology/2009/apr/30/e–votingelectronic–polling–systems, свободный.

18. Baudron, O. Practical multi–candidate election system. In proceedings of the twentieth annual ACM symposium on Principles of distributed computing, // Baudron, O., Fouque, P.A., Pointcheval, D., Stern, J., Poupard, G. – ACM, 2001. –pp. 274 – 283.

References

1. Prasti N. Blokcheyn. Razrabotka prilozheniy, // N. Prasti, V.S. Yatsenkov. – SPb.: BKHV–Peterburg, 2018. – 256 s.

2. Raval S. Detsentralizovannyye prilozheniya. Tekhnologiya Blockchain v deystvii, // S. Raval. – SPb.: Piter, – 2017. – 192 s.

3. Tapskott D., Tapskott A. Tekhnologiya blokcheyn – to, chto dvizhet finansovoy revolyutsiyey segodnya, // D. Tapskott, A. Tapskott. – M.: Eksmo, 2017. – 448 s.

4. Naskol'ko nadezhno elektronnoye golosovaniye [Elektronnyy resurs]. – Rezhim dostupa: https://www.svoboda.org/a/269300.html, svobodnyy.

5. Norvegiya ofitsial'no otkazalas' ot elektronnogo golosovaniya na vyborakh: ono kontrproduktivno [Elektronnyy resurs]. – Rezhim dostupa: http://www.mk.ru/politics/world/2014/06/30/norvegiya–otkazalasv–politike–ot–elektronnogo–golosovaniya.html, svobodnyy.

6. Block The Vote: Could Blockchain Technology Cybersecure Elections? [Elektronnyy resurs]. – Rezhim dostupa: http://www.forbes.com/sites/realspin/2016/08/30/block–the–votecouldblockchain–technology–cybersecure–elections, svobodnyy.

7. California: The Top to Bottom Review [Electronic resource]. – Access mode: http://www.votetrustusa.org/index.php?option=com_content&task=view&id =2554&Itemid=113, free.

8. IGS Votomatic Prototype Goes to the Smithsonian [Electronic resource]. – Access mode: https://web.archive.org/web/20070713201451/http://www.igs.berkeley.edu/publications/par/winter2001/votomatic.htm, free.

9. Kiwi. Bitcoin testnet sandbox. [Electronic resource]. – Access mode: https://testnet.manu.backend.hamburg/faucet, free.

10. NSW election result could be challenged over iVote security flaw [Electronic resource]. – Access mode: https://www.theguardian.com/australia–news/2015/mar/23/nsw–electionresult–could–be–challenged–over–ivote–security–flaw, free.

11. Peer–to–peer [Electronic resource]. – Access mode: https://bitcoin.org/bitcoin.pdf, free.

12. Russian Hackers Acted to Aid Trump in Election, U.S. Says [Electronic resource]. – Access mode: https://www.nytimes.com/2016/12/09/us/obama–russia–election–hack.html, free.

13. Slim. Middeware–slim. [Electronic resource]. – Access mode: https://www.slimframework.com/docs/v3/concepts/middleware.html, free.

14. State bans electronic balloting in 4 counties / Touch–screen firm accused of 'reprehensible,' illegal conduct [Electronic resource]. – Access mode: https://www.sfgate.com/politics/article/State–bans–electronic–balloting–in–4– counties–2784975.php, free.

15. Top 100 Cryptocurrencies by Market Capitalization [Электронный ресурс]. – Режим доступа: https://coinmarketcap.com/, свободный.

16. Voting Machine Company Submits to Inquiry [Электронный ресурс]. – Режим доступа: https://www.nytimes.com/2006/10/31/us/politics/31vote.html?_r=1&oref=slogin, свободный.

17. Why machines are bad at counting votes [Электронный ресурс]. – Режим доступа: https://www.theguardian.com/technology/2009/apr/30/e–votingelectronic–polling–systems, свободный.

18. Baudron, O. Practical multi–candidate election system. In proceedings of the twentieth annual ACM symposium on Principles of distributed computing, // Baudron, O., Fouque, P.A., Pointcheval, D., Stern, J., Poupard, G. – ACM, 2001. –pp. 274 – 283.

Кротова Елена Львовна, кандидат физико-математических наук, доцент кафедры «Высшая математика», Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: lenkakrotova@yandex.ru

Субботина Юлия Владимировна, ведущий инженер кафедры «Высшая математика», аспирант кафедры «Высшая математика», Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: yulyia.urazbaeva@mail.ru

Ермаков Дмитрий Германович, кандидат физико-математических наук, старший научный сотрудник отдела дифференциальных уравнений Лаборатории научно-информационных ресурсов, Федеральное государственное бюджетное учреждение науки Институт математики и механики им. Н. Н. Красовского Уральского отделения Российской академии наук (ИММ УрО РАН). 620108, г. Екатеринбург, ул. Софьи Ковалевской, д. 16.; кандидат физико-математических наук, доцент, Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина». 620002, Екатеринбург, ул. Мира, 19. E-mail: Dmitry.Ermakov@mail.ru

Тишин Константин Львович, аспирант кафедры «Высшая математика», Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: konstantinlvovich777@gmail.com

Krotova Elena Lvovna, Candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Higher Mathematics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: lenkakrotova@yandex.ru

Subbotina Yulia Vladimirovna, Leading engineer of the Department of Higher Mathematics, postgraduate student of the Department of Higher Mathematics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: yulyia.urazbaeva@mail.ru

Yermakov Dmitry Germanovich, Candidate of Physical and Mathematical Sciences, Senior Researcher, Department of Differential Equations, Laboratory of Scientific and Information Resources, Federal State Budgetary Scientific Institution, N. N. Krasovsky Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences (IMM UB RAS). 620108, Yekaterinburg, Sofya Kovalevskaya St., 16; Candidate of Physical and Mathematical Sciences, Associate Professor, Federal State Autonomous Educational Institution of Higher Education "Ural Federal University named after the first President of Russia B. N. Yeltsin". 620002, Yekaterinburg, Mira St., 19. E-mail: Dmitry.Ermakov@mail.ru.

Tishin Konstantin Lvovich, postgraduate student, Department of Higher Mathematics, Perm National Research Polytechnic University. 614990, Perm Krai, Perm, Komsomolsky Prospekt, 29. E-mail: konstantinlvovich777@gmail.com.

УДК 004.056 Вестник УрФО № 4(54) / 2024, с. 22–29

Белоногов А. С., Будник М. Г., Мельников А. В. DOI: 10.14529/secur240403

ИНСТРУМЕНТАЛЬНАЯ СРЕДА РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В статье исследовано взаимодействие открытой системы IRP, разрабатываемой международным сообществом, с межсетевыми экранами российского производства. На основе функций системы «TheHive» реализован функционал по блокированию атак с внешнего периметра. Представлен алгоритм модуля взаимодействия с межсетевыми экранами для предотвращения компьютерных атак, для разработки которого использован язык Python, в частности, модуль textFSM. Результатом является IRP-система с реализованной пользовательской функцией реагирования, которая может послужить технологической основой в работе Центра обеспечения безопасности (SOC).

Ключевые слова: *IRP-система, SOC, информационная безопасность, кибератака, реагирование на инциденты, центр обеспечения безоапасности.*

Belonogov A. S., Budnik M. G., Melnikov A. V.

INFORMATION SECURITY INCIDENT RESPONSE TOOLING ENVIRONMENT

The article examines the interaction of an open IRP system developed by the international community with Russian-made firewalls. Based on the functions of the "TheHive" system, the functionality for blocking attacks from the external perimeter is implemented. The algorithm of the module for interaction with firewalls to prevent computer attacks is presented, for the development of which the Python language is used, in particular, the textFSM module. The result is an IRP system with an implemented user response function, which can serve as a technological basis for the work of the Security Center (SOC).

Keywords: *IRP system, SOC, information security, cyberattack, incident response, security operations center.*

Введение.

Киберпреступления ежегодно приносят серьезный урон государственным и коммерческим организациям. Согласно [1] за 2023 год ущерб от IT-преступлений превысил в РФ 156 млрд. руб., в то же время по оценке [2] в США заявленный ущерб за тот же период составил 12,5 млрд. долларов США. С ростом объема и ценности обрабатываемых данных соответствующим образом должны развиваться и способы защиты информации от кибератак. Обеспечение информационной безопасности – это непрерывный процесс, задача с неснижающейся актуальностью. Чтобы решать ее с достаточным уровнем оперативности и качества при определенных размерах компании целесообразным может стать создание ведомственного или привлечение стороннего Центра обеспечения безопасности (Security Operation Center, SOC). В компании такой центр обеспечивает непрерывную защиту организации от киберугроз. Центр обеспечения безопасности в своей деятельности опирается на различные технические средства (рис.1), обеспечивающие его функциональность, однако обязательными можно считать базовые средства мониторинга и реагирования. [3, 4]

При этом отмечается постоянно растущая нагрузка на [5,6] специалистов SOC и рутинный характер работы [7], приводящий, в том числе, к высокой текучести кадров. [4] Попытки облегчить и ускорить рутинную работу аналитика SOC проводятся в [5-10], преимущественно с применением технологий искусственного интеллекта. Делать это предполагается, например, за счет автоматизации подготовки плейбуков реагирования на кибератаки [6], или созданию помощника с искусственным интеллектом [5], и прочими способами. В данной работе мы рассмотрим автоматизацию одного из рутинных действий аналитика SOC с применением системы класса IRP.

Системы IRP.

Для автоматизации реагирования на инциденты, совместной работы над расследованием и реагированием на инциденты, построения отчетности применяются системы класса IRP (Incident Response Platform). Данный класс систем позволяет повысить эффективность про-цессов работы над инцидентом. [11]

Перечислим базовые функции IRP-систем:

- регистрация инцидента;
- совместная работа специалистов при реагировании на инцидент;
- эскалация и оповещение;
- интеграция с существующими в компании средствами;
- автоматизированное реагирование на инциденты;
- база знаний;
- отчеты. [12]

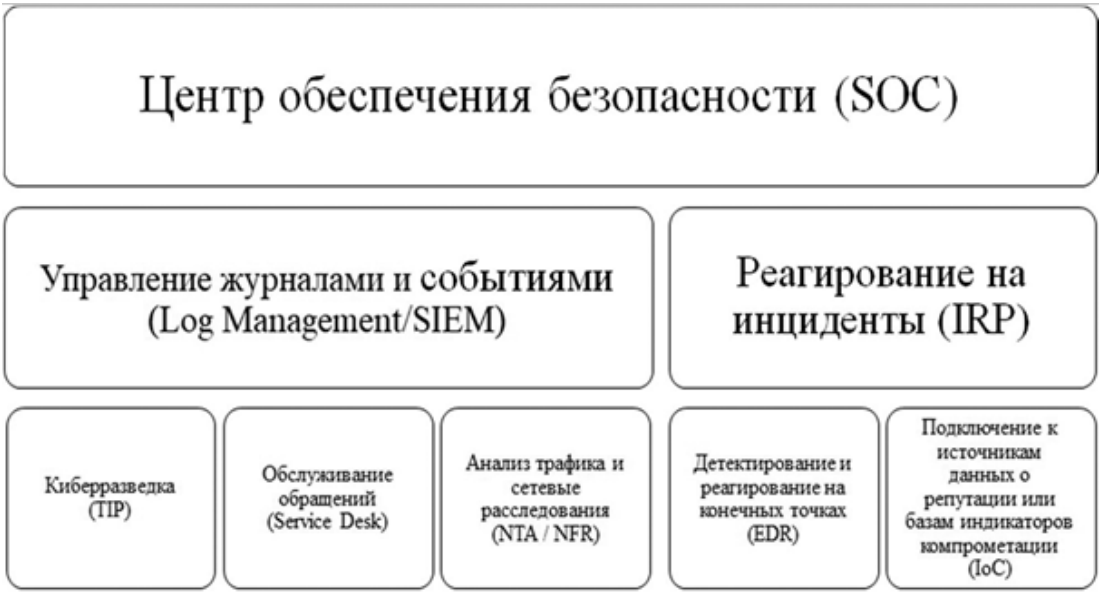


Рис. 1. Функциональная структура центра обеспечения безопасности

Таблица 1.

Заявленные интеграции российских производителей IRP-систем с межсетевыми экранами

	Security Vision [13]	**R-Vision SOAR [14]**	**Innostage Orchestrator [15]**
Зарубежные производители межсетевых экранов	Cisco ASA Fortigate Check Point Cisco Firepower Juniper	Не заявлено	Check Point Palo Alto Huawei
Российские производители межсетевых экранов	Не заявлено	Не заявлено	Usergate

Для реализации автоматизированного реагирования на инциденты IRP система должна иметь возможность интегрироваться с существующими в организации средствами защиты информации. При этом обычной ситуацией является использование для целей обеспечения защиты информации нескольких типов продуктов и инструментов от различных производителей. Современные продукты могут работать независимо, и, как правило, имеют собственные механизмы представления данных, не соблюдающие стандартизацию для обмена данными. [10]

Так, например, коммерческие IRP от российских производителей заявляют поддержку российских средств защиты информации типа межсетевой экран в довольно ограниченном объеме, что отображено в Таблице 1.

Система с открытым исходным кодом TheHive во взаимодействии с продуктом той же команды разработчиков Cortex позволяет автоматизировать запуск определенных функций. В системе TheHive этот процесс реализован через механизм запуска анализаторов (Analyzer) и ответчиков (Responder). Подобный функционал может быть использован командой SOC для прерывания выявленной вредоносной деятельности путем блокирования сетевого взаимодействия с указанным внешним IP-адресом на пограничном межсетевом экране.

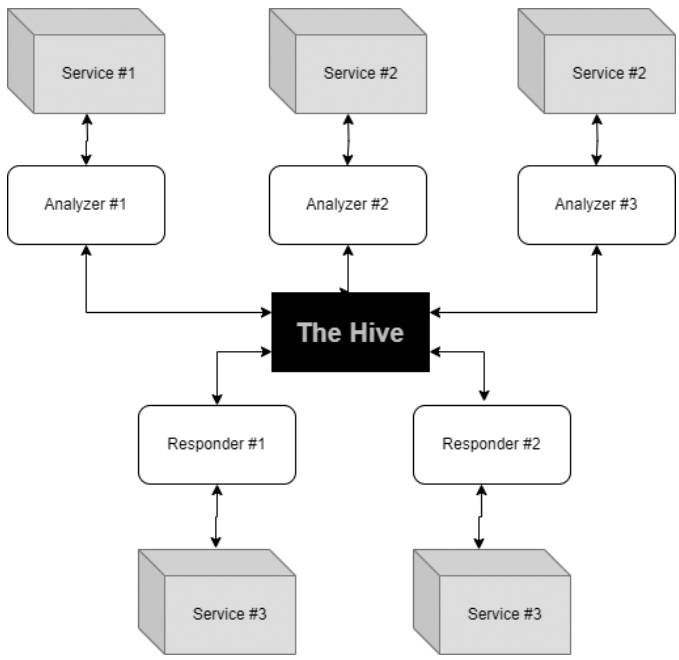


Рис. 2. Структурная схема TheHive.

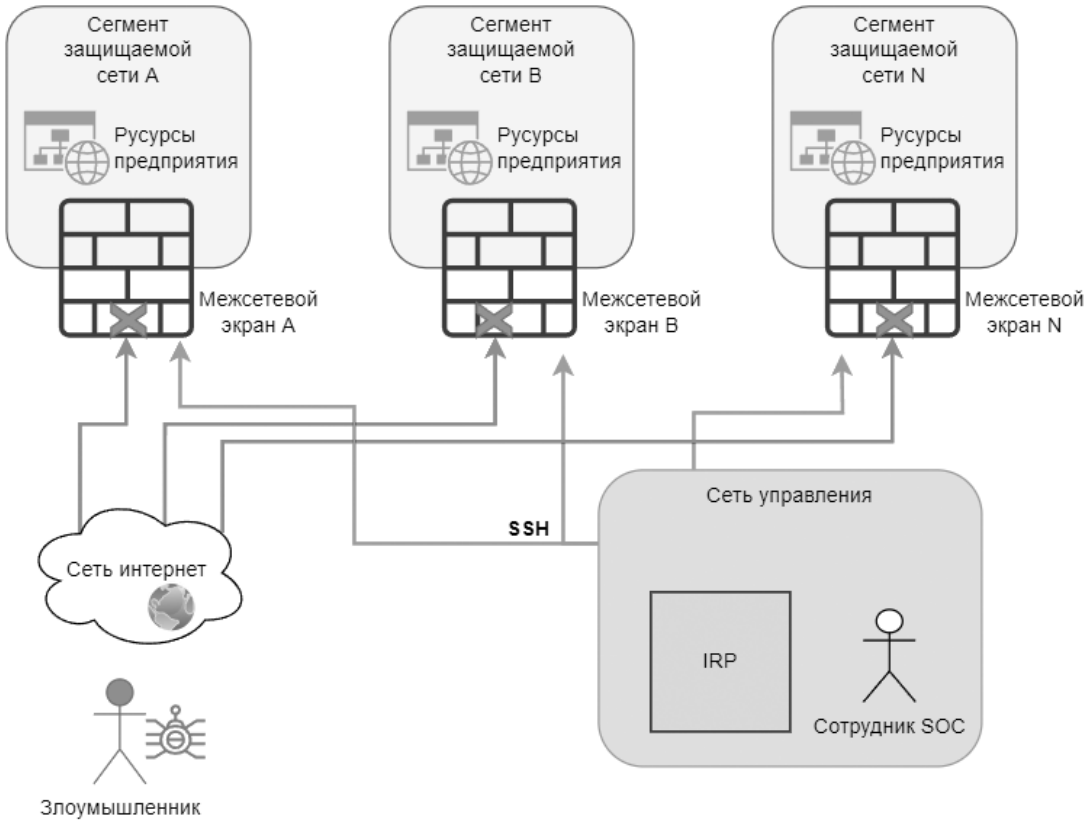


Рис. 3. Общая схема сетевых воздействий на межсетевые экраны при входящей атаке на ресурсы

Разработчики системы TheHive опубликовали руководство для создания собственных функций реагирования (ответчиков), что позволяет самостоятельно продумать и реализовать возможные способы решения задач защиты в конкретной инфраструктуре (рис. 2). [16].

Реализация блокировки атакующих IP-адресов.

В рамках рассматриваемого примера необходимо упомянуть об особенностях атак с внешнего периметра. Как правило наиболее типовыми будут: сканирование с целью выявления открытых сервисов инфраструктуры (разведка), использование в автоматическом режиме скриптов, экс-плуатирующих известные уязвимости, например, SQL-инъекции.

Мерой противодействия таким атакам является добавление правила на пограничный межсетевой экран для блокировки трафика от IP-адреса нарушителя. Необходимость автоматизации таких действий рассматривается как одна из первоочередных в [6]. Потребность усиливается при наличии в обслуживании SOC нескольких сегментов защищаемых компьютерных сетей предприятий.

Рассмотрим блокирование атаки на межсетевом экране российского производства ПАК ViPNet xFirewall 5. Через использование функционала ответчика (Responder) эксперт центра обеспечения безопасности может моментально отправить данные атакующего узла для создания блокирую-щего правила на межсетевом экране (рис. 3).

Для конфигурации функции у пользователя необходимо затребовать следующие параметры:

- IP-адрес управления межсетевым экраном;
- порт управления межсетевого экрана;
- пользователь межсетевого экрана;
- пароль пользователя межсетевого экрана;
- пароль перехода в привилегированный режим.

Для создания ответчика с применением языка Python необходимо подготовить 3 файла: файл с кодом алгоритма на языке Python, файл с перечислением всех зависи-

мостей, используемых в коде программы, и файл взаимодействия со службой, описывающий в формате JSON ключевые параметры взаимодействия с функцией.

Приведем примеры параметров, обязательных для заполнения в вышеуказанном JSON-файле:

• DataTypeList – список типов данных TheHive, поддерживаемых ответчиком;

• Command – это относительный путь к исполняемому файлу програм-мы, в нашем случае: IPBlock/ipblock.py;

• ConfigurationItems – список элементов конфигурации, предназначенных для установки всех переменных ответчиков непосредственно в пользовательском интерфейсе Cortex.

Особенностью рассматриваемого межсетевого экрана является отсутствие опубликованных API для взаимодействия с его конфигурацией. Вследствие чего работа пользовательского кода основана на подключении к консоли управления ПАК по протоколу SSH. Дальнейшая работа автоматической конфигурации производится посредством анализа структурированного вывода (рис. 4). С целью проверки результатов выполнения функции Responder весь вывод, полученный от межсетевого экрана анализируется при помощи модуля textFSM. [17]

Заключение

Приведенная реализация сервиса блокирования целена-правленных атак на инфраструктуру извне расширила возможности системы TheHive, позволяя увеличить скорость реакции команды Центра обеспечения безопасности на киберугрозы. Благодаря принципу открытости исследуемой IRP-системы, становится возможно произвести настройку и интеграцию практически в любой конфигурации компьютерной сети, что, в свою очередь, является серьезным подспорьем для создания ведомственного Центра обеспечения безопасности без высоких первоначальных затрат. Для повышения качества и эффективности реагирования на киберугрозы актуальной задачей является разработка ответчиков (Responders), способных взаимодействовать со средствами технической защиты информации российского производства.

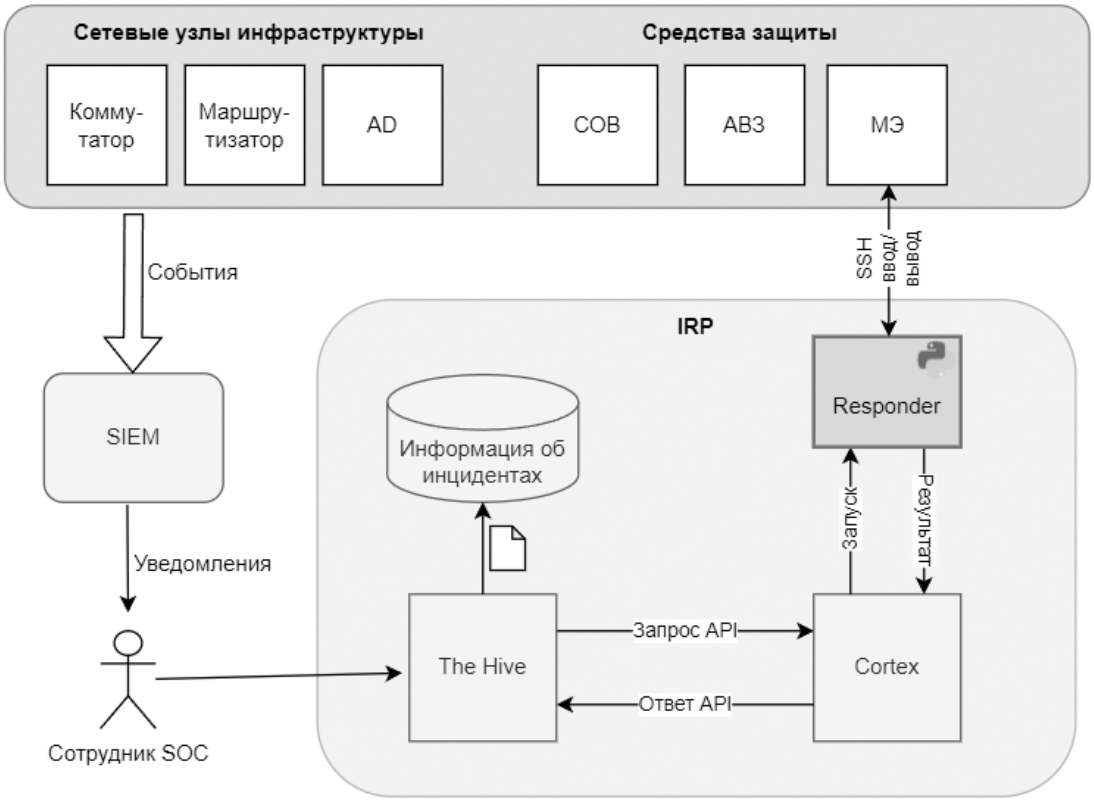


Рис. 4. Схема работы сотрудника SOC с применением TheHive и разработанного ответчика

Литература

1. Расширенное заседание коллегии МВД 2 апреля 2024 года, 15:00 Москва. Стенограмма доклада В.В. Путина. Режим доступа: http://kremlin.ru/events/president/news/73770 (дата обращения 07.11.2024 г.)

2. I.C.C. Center. Internet Crime Complaint Center. Режим доступа: https://www.ic3.gov/ (дата обращения 01.11.2024 г.)

3. Очередько А.Р. Исследование IRP-систем на основе анализа механизмов реагирования на инциденты информационной безопасности / А. Р. Очередько, Д.А. Бачманов, М.М. Путято, А. С. Макарян // Прикаспийский журнал: управление и высокие технологии. – 2021. – № 1. – С. 74–82.

4. Manfred Vielberth, Fabian Böhm, Ines Fichtinger, Günther Pernul Se-curity Operations Center: A Systematic Study and Open Challenges // IEEE Access (Volume: 8), P. 227756–227779 – 2020 – DOI: 10.1109/ACCESS.2020.3045514 – Режим доступа: https://ieeexplore.ieee.org/document/9296846

5. Scott Freitas, Amir Gharib, Jovan Kalajdjieski, Robert McCann AI-Driven Guided Response for Security Operation Centers with Microsoft Copilot for Security – 2024 – DOI: 10.48550/arXiv.2407.09017 – Режим доступа: https://arxiv.org/abs/2407.09017

6. Ryuta Kremer, Prasanna N. Wudali, Yuval Elovici, Asaf Shabtai, Satoru Momiyama, Toshinori Araki, Jun Furukawa, IC-SECURE: Intelligent System for Assisting Security Experts in Generating Playbooks for Automated Incident Response – DOI: 10.48550/arXiv.2311.03825 – Режим доступа: https://arxiv.org/pdf/2311.03825

7. Andreas U. Schmidt, Sven Knudsen, Tobias Niehoff, and Klaus Schwietz Planning Distributed Security Operations Centers in Multi-Cloud Landscapes: A Case Study – 2023 – DOI: 10.48550/arXiv.2303.03141 – Режим доступа: https://arxiv.org/abs/2303.03141

8. PeiYu Tseng, ZihDwo Yeh, Xushu Dai, Peng Liu Using LLMs to Au-tomate Threat Intelligence Analysis Workflows in Security Operation Centers // JOURNAL OF LATEX CLASS FILES, Vol. 18, No. 9 – 2020 – DOI: 10.48550/arXiv.2407.13093 – Режим доступа: https://arxiv.org/pdf/2407.13093

9. Hari Hayagreevan, Souvik Khamaru Security of and by Generative AI platforms // Whitepaper February 2024 – DOI: 10.48550/arXiv.2410.13899 – Режим доступа: https://arxiv.org/pdf/2410.13899

10. Johnson Kinyua, Lawrence Awuah AI/ML in Security Orchestration, Automation and Response: Future Research Directions // Intelligent Automation & Soft Computing 2021 Vol.28, No.2, P. 527-545 – 2021 – DOI: 10.32604/iasc.2021.016240 – Режим доступа: https://www.techscience.com/iasc/v28n2/42057/pdf

11. Обзор рынка платформ реагирования на инциденты (IRP) в России. – 2018. – Режим доступа: https://www.anti-malware.ru/analytics/Market_Analysis/incident-response-platforms-irp-in-russia, свободный. – Заглавие с экрана. – Яз. рус. (дата обращения: 26.10.2023 г.)

12. Бесплатная IRP-система своими силами: опыт использования платформы с открытым кодом The Hive. – 2019. – Режим доступа: https://www.anti-malware.ru/practice/solutions/free-IRP-on-your-own, свобод-ный. – Заглавие с экрана. – Яз. рус. (дата обращения: 27.10.2023 г.)

13. Опросный лист Security Vision для подготовки ТКП. Режим доступа: https://www.securityvision.ru/downloads/%D0%9E%D0%BF%D1%80%D0%BE%D1%81%D0%BD%D1%8B%D0%B9%20%D0%BB%D0%B8%D1%81%D1%82%20Security%20Vision.xlsx (дата обращения 21.06.2024)

14. Опросный лист R-Vision SOAR. Режим доступа: https://rvision.ru/blog-posts/usloviya-priobreteniya-po (дата обращения 21.06.2024)

15. Innostage Orchestrator Система управляющих воздействий. Режим доступа: https://inno-orch.ru/ – Заглавие с экрана. (дата обращения 24.06.2024)

16. The Hive-Project. – 2020. – Режим доступа: https://github.com/TheHive-Project/TheHive, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 24.10.2023 г.)

17. TextFSM. – 2023. – Режим доступа: https://github.com/google/textfsm/wiki/TextFSM, свободный. – Заглавие с экрана. – Яз. англ. (дата обращения: 26.10.2023 г.).

References

1. Rasshirennoye zasedaniye kollegii MVD 2 aprelya 2024 goda, 15:00 Moskva. Stenogramma doklada V.V. Putina. Rezhim dostupa: http://kremlin.ru/events/president/news/73770 (data obrashcheniya 07.11.2024 g.)

2. I.C.C. Center. Internet Crime Complaint Center. Rezhim dostu-pa: https://www.ic3.gov/ (data obrashcheniya 01.11.2024 g.)

3. Ochered'ko A.R. Issledovaniye IRP-sistem na osnove analiza mekha-nizmov reagirovaniya na intsidenty informatsionnoy bezopasnosti / A. R. Ochered'ko, D.A. Bachmanov, M.M. Putyato, A. S. Makaryan // Pri-kaspiyskiy zhurnal: upravleniye i vysokiye tekhnologii. – 2021. – № 1. – S. 74–82.

4. Manfred Vielberth, Fabian Böhm, Ines Fichtinger, Günther Pernul Se-curity Operations Center: A Systematic Study and Open Challenges // IEEE Access (Volume: 8), P. 227756–227779 – 2020 – DOI: 10.1109/ACCESS.2020.3045514 – Rezhim dostupa: https://ieeexplore.ieee.org/document/9296846

5. Scott Freitas, Amir Gharib, Jovan Kalajdjieski, Robert McCann AI-Driven Guided Response for Security Operation Centers with Microsoft Copilot for Security – 2024 – DOI: 10.48550/arXiv.2407.09017 – Rezhim dostupa: https://arxiv.org/abs/2407.09017

6. Ryuta Kremer, Prasanna N. Wudali, Yuval Elovici, Asaf Shabtai, Satoru Momiyama, Toshinori Araki, Jun Furukawa, IC-SECURE: Intelligent System for Assisting Security Experts in Generating Playbooks for Automated Incident Response – DOI: 10.48550/arXiv.2311.03825 – Rezhim dostupa: https://arxiv.org/pdf/2311.03825

7. Andreas U. Schmidt, Sven Knudsen, Tobias Niehoff, and Klaus Schwietz Planning Distributed Security Operations Centers in Multi-Cloud Landscapes: A Case Study – 2023 – DOI: 10.48550/arXiv.2303.03141 – Rez-him dostupa: https://arxiv.org/abs/2303.03141

8. PeiYu Tseng, ZihDwo Yeh, Xushu Dai, Peng Liu Using LLMs to Au-tomate Threat Intelligence Analysis Workflows in Security Operation Centers // JOURNAL OF LATEX CLASS FILES, Vol. 18, No. 9 – 2020 – DOI: 10.48550/arXiv.2407.13093 – Rezhim dostupa: https://arxiv.org/pdf/2407.13093

9. Hari Hayagreevan, Souvik Khamaru Security of and by Generative AI platforms // Whitepaper February 2024 – DOI: 10.48550/arXiv.2410.13899 – Rezhim dostupa: https://arxiv.org/pdf/2410.13899

10. Johnson Kinyua, Lawrence Awuah AI/ML in Security Orchestration, Automation and Response: Future Research Directions // Intelligent Automation & Soft Computing 2021 Vol.28, No.2, P. 527-545 – 2021 – DOI: 10.32604/iasc.2021.016240 – Rezhim dostupa: https://www.techscience.com/iasc/v28n2/42057/pdf

11. Obzor rynka platform reagirovaniya na intsidenty (IRP) v Rossii. – 2018. – Rezhim dostupa: https://www.anti-malware.ru/analytics/Market_Analysis/incident-response-platforms-irp-in-russia, svobodnyy. – Zaglaviye s ekrana. – Yaz. rus. (data obrashcheniya: 26.10.2023 g.)

12. Besplatnaya IRP-sistema svoimi silami: opyt ispol'zovaniya platformy s otkrytym kodom The Hive. – 2019. – Rezhim dostupa: https://www.anti-malware.ru/practice/solutions/free-IRP-on-your-own, svobodnyy. – Zaglaviye s ekrana. – Yaz. rus. (data obrashcheniya: 27.10.2023 g.)

13. Oprosnyy list Security Vision dlya podgotovki TKP. Rezhim dostupa: https://www.securityvision.ru/downloads/%D0%9E%D0%BF%D1%80%D0%BE%D1%81%D0%BD%D1%8B%D0%B9%20%D0%BB%D0%B8%D1%81%D1%82%20Security%20Vision.xlsx (data obrashcheniya 21.06.2024)

14. Oprosnyy list R-Vision SOAR. Rezhim dostupa: https://rvision.ru/blog-posts/usloviya-priobreteniya-po (data obrashcheniya 21.06.2024)

15. Innostage Orchestrator Sistema upravlyayushchikh vozdeystviy. Rez-him dostupa: https://innoorch.ru/ – Zaglaviye s ekrana. (data obrashche-niya 24.06.2024)

16. The Hive-Project. – 2020. – Rezhim dostupa: https://github.com/TheHive-Project/TheHive, svobodnyy. – Zaglaviye s ekrana. – Yaz. angl. (data obrashcheniya: 24.10.2023 g.)

17. TextFSM. – 2023. – Rezhim dostupa: https://github.com/google/textfsm/wiki/TextFSM, svobodnyy. – Zaglaviye s ekrana. – Yaz. angl. (data obrashcheniya: 26.10.2023 g.).

Белоногов Александр Сергеевич, руководитель центра сетевых технологий и телекоммуникаций, автономное учреждение Ханты-Мансийского автономного округа – Югры «Югорский научно-исследовательский институт информационных технологий». 628011, г. Ханты-Мансийск, ул. Мира, 151. E-mail: BelonogovAS@uriit.ru

Будник Максим Геннадьевич, начальник отдела развития и автоматизации ИТ инфраструктуры центра сетевых технологий и телекоммуникаций, автономное учреждение Ханты-Мансийского автономного округа – Югры «Югорский научно-исследовательский институт информационных технологий». 628011, г. Ханты-Мансийск, ул. Мира, 151. E-mail: BudnikMG@uriit.ru

Мельников Андрей Витальевич, доктор технических наук, профессор, директор, автономное учреждение Ханты-Мансийского автономного округа – Югры «Югорский научно-исследовательский институт информа-ционных технологий». 628011, г. Ханты-Мансийск, ул. Мира, 151. E-mail: MelnikovAV@uriit.ru

Belonogov Alexander Sergeevich, Head of the Center for Network Technologies and Telecommunications, Ugra Research Institute of Information Technologies. 628011, Khanty-Mansiysk, Mira str., 151. Email: Belonogo-vAS@uriit.ru

Budnik Maxim Gennadievich, Head of the IT Infrastructure Development and Automation Department of the Center for Network Technologies and Telecommunications, Ugra Research Institute of Information Technologies. 628011, Khanty-Mansiysk, Mira str., 151. Email: BudnikMG@uriit.ru

Melnikov Andrey Vitalievich, Doctor of Technical Sciences, Professor, Director, Ugra Research Institute of Information Technologies. 628011, Khanty-Mansiysk, Mira str., 151. Email: MelnikovAV@uriit.ru

УДК 004.774 Вестник УрФО № 4(54) / 2024, с. 30–36

Жусов Д. Л., Макеев С. М., Соколов А. Н. DOI: 10.14529/secur240404

АНАЛИЗ СТРУКТУРЫ WEB-САЙТОВ ДЛЯ ИДЕНТИФИКАЦИИ ОБЪЕКТОВ ЭКСПЛУАТАЦИИ УГРОЗ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В работе представлен вариант решения задачи определения потенциальных угроз информационной безопасности web-сайтов через анализ их структуры и сбор статистических данных о посещаемости отдельных информационных ресурсов. Представлено алгоритмическое обеспечение рассмотренной процедуры. Описана структура данных, обобщающих сведения для идентификации угроз web-сайтов.

Ключевые слова: *web-сайт, гиперссылка, гипертекстовый переход, структура web-сайта, информационная безопасность*

Zhusov D. L., Makeev S. M., Sokolov A. N.

ANALYSIS OF THE STRUCTURE OF WEB SITES TO IDENTIFY OBJECTS OF EXPLOITATION OF INFORMATION SECURITY THREATS

The paper presents a solution to the problem of identifying potential threats to the information security of web sites through the analysis of their structure and the collection of statistical data on the attendance of individual information resources. The algorithmic support of the considered procedure is presented. The structure of data summarizing information for the identification of threats to web sites is described.

Keywords: *website, hyperlink, hypertext navigation, website structure, information security*

Введение

Совершенствование механизмов государственного управления обусловило широкое применение современных информационных технологий для организации взаимодействия государства и общества. Важная роль в структуре инструментов совершенствования информационного общества страны отводится web-сайтам, функционирующих в сети Интернет и обеспечивающих доступ к информации для широкого круга пользователей [1]. Вопросы эффективного функционирования таких инструментов с функциональной точки зрения нормативно определены в [1, 2], а с точки зрения информационной безопасности – в [3].

Компоненты web-сайта в форме исполняемых модулей (сценариев), предназначенных для динамического формирования информационных ресурсов, могут выступать в качестве источников реализации различных классов компьютерных атак [4]. Статистика Positive Technologies [5] свидетельствует о росте числа компьютерных атак на web-приложения. Применение межсетевых экранов web-приложений направлено на обнаружение и блокировку сетевых компьютерных атак, однако не предоставляет возможности анализа инструментария их реализации – конкретного скрипта (приложения), в составе которого присутствуют фрагменты кода – угроз информационной безопасности. Таким образом, можно предположить, что обеспечение информационной безопасности web-сайтов может быть напрямую связано с идентификацией их компонентов, которые потенциально могут быть использованы нарушителем для реализации компьютерных атак. Учитывая, что в качестве таких компонентов web-сайтов выступают информационные ресурсы (web-приложения), то актуальной является задача определения их структуры и последующей оценки количества обращений к ним с привязкой посещаемости к потенциальной эксплуатации уязвимости. Следовательно, анализ структуры web-сайтов может способствовать идентификации источников реализации угроз информационной безопасности и будет в конечном итоге определять повышение защищенности web-сайтов.

Результаты мониторинга [6] свидетельствуют об отсутствии единого подхода к формированию структуры web-сайтов, а существующие способы проверки web-сайтов зачастую слабо автоматизированы, что делает процесс анализа достаточно трудоемким и затратным по времени.

Исследования [7] свидетельствуют о возможности представления структуры web-сайта (рис. 1) графом $G=(V,E)$, где V – множество вершин графа (web-страниц), а E – множество ребер графа (гипертекстовых переходов по сайту – гиперссылок). Для расчета глубины гипертекстовых переходов могут быть использованы известные алгоритмы обхода вершин графа – «*поиск в глубину*» и «*поиск в ширину*» [7].

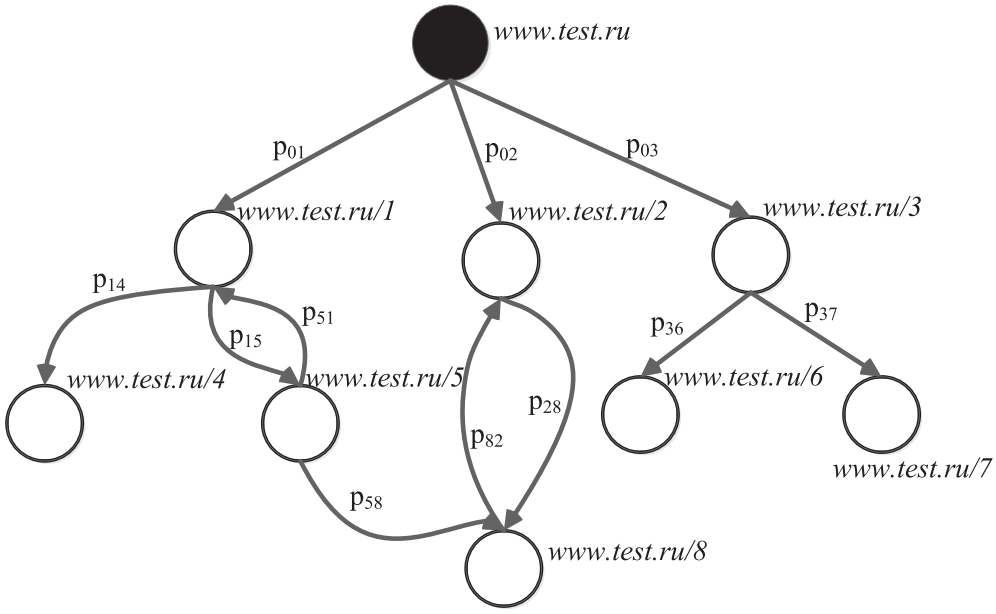


Рис. 1. Графовое представление web-сайта

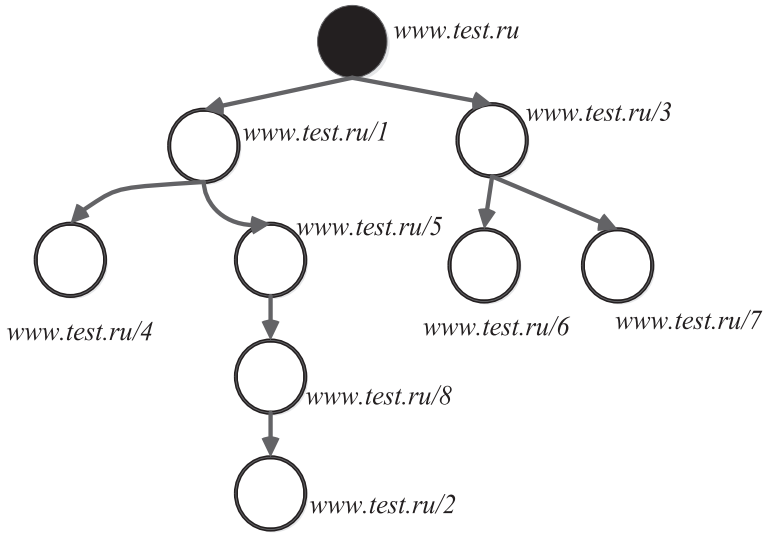


Рис. 2. Дерево web-сайта, построенное алгоритмом поиска «в глубину»

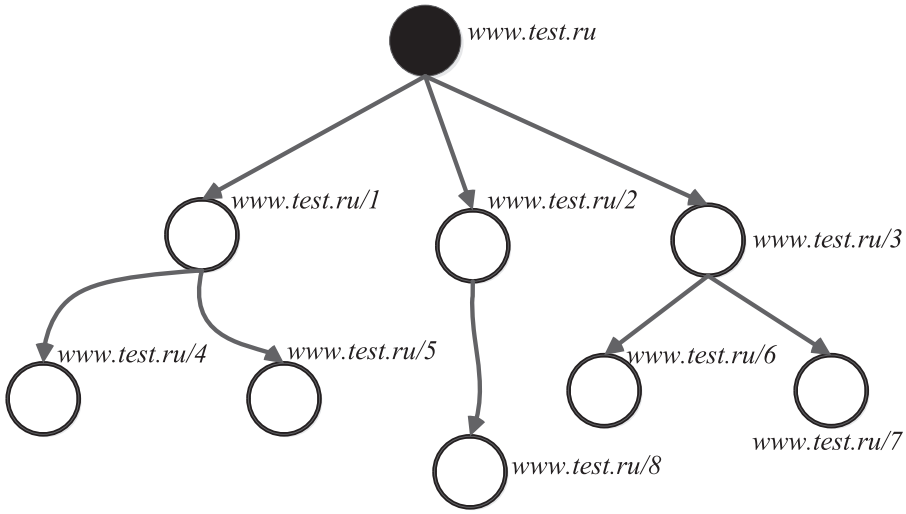


Рис. 3. Дерево web-сайта, построенное алгоритмом поиска «в ширину»

Дерево web-сайта, построенное по алгоритму «в глубину» (рис. 2), и рассчитанное значение максимальной глубины гипертекстовых переходов (равное 4-м) не соответствует действительности, что позволяет говорить о невозможности применения данного алгоритма. В то же время, дерево web-сайта, построенное по алгоритму «в ширину» (рис. 3), и значение максимальной глубины гипертекстовых переходов (равное 3-м) соответствует истинной структуре web-сайта (рис. 1).

Рассчитанная в результате поиска «в ширину» максимальная глубина гипертекстовых переходов равна 3, что соответствует фактическому значению (рис. 1) и свидетельствует о целесообразности применения алгоритма поиска «в ширину» для решения поставленной задачи.

Проведенные исследования позволили разработать алгоритм анализа структуры web-сайтов (рис. 4).

Основу разработанного алгоритма составляет набор рекурсивных процедур, представленный блоками 3 – 15. Сложность его определяется поиском в коде загруженной web-страницы тегов $<a>$ с атрибутом *href* и количеством страниц анализируемого web-сайта и вычисляется как

$$O(\eta) \cdot O(V + E),$$

где η – количество символов в коде web-страниц.

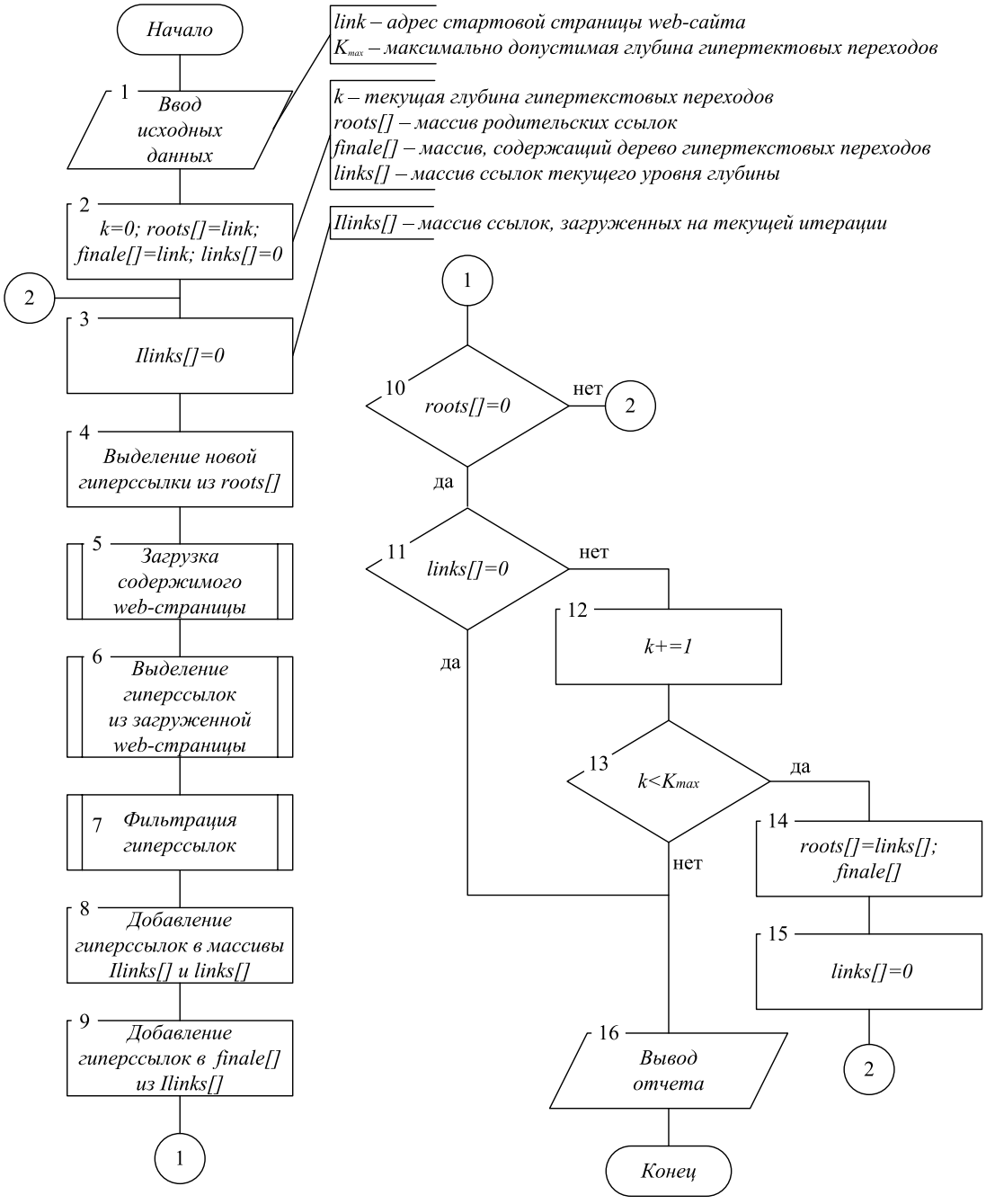


Рис. 4. Алгоритм анализа структуры web-сайта

Каждому узлу сформированного дерева web-сайта ставится в соответствие адрес доступа к нему. Это позволяет сформировать в табличной форме для каждого модуля обработки (сценария, см. табл. 1) не только взаимосвязи с другими информационными ресурсами, но и с полученными данными интегрированной в web-сайт системы статистики.

Статистический анализ данных количества посещений web-сайта позволит определить либо значимые события, вызвавшие интерес со стороны пользователей web-сайта, либо необходимость более углубленного анализа данных при отсутствии публикаций за конкретный период.

Таблица 1

Фрагмент данных анализа структуры web-сайта

Адрес ресурса	Количество взаимосвязей	Глубина переходов	Количество посещений
http://test.ru/index.html	15	0	100
http://test.ru/1/scr1.php	4	1	250
http://test.ru/1/1/scr2.php	2	2	3500
http://test.ru/1/2/scr3.php	2	2	100

Рассмотрим потенциальную ситуацию на примере (табл. 1).

В представленном примере (табл. 1) иллюстрируется ситуация, при которой количество посещений стартовой страницы (100) равно количеству обращений к сценарию *scr3*(100), однако это существенно меньше числа обращений к другим компонентам web-сайта – сценариям *scr1*(250) и *scr2* (3500). Такие сценарии являются серверным расширением функционала web-сайта и могут являться потенциальными объектами эксплуатации угроз информационной безопасности.

С точки зрения реализации угроз функционированию web-сайта (в первую очередь контроля целостности и доступности web-сайта) это означает, что серверные сценарии эксплуатируются (выполняются) существенно чаще числа обращений к стартовой странице. Для исполнения сценариев необходимы значения, которые вводятся пользователем в соответствующих формах на страницах web-сайта, которые являются значениями параметров, передаваемых сценариям. Такой набор статистических данных при известной структуре web-сайта может являться потенциально опасной ситуацией. Это требует более детального анализа сетевого трафика к защищаемым ресурсам или анализа исходного кода серверных сценариев на предмет эксплуатации уязвимостей межсайтового выполнения сценариев, SQL-инъекций и т.п. [8].

Возможным направлением дальнейших исследований может являться углубленный статистический анализ данных о посещаемости информационных ресурсов web-сайтов (обращении к ним) в зависимости от источников и пролонгированных во времени наборах обращений к web-сайтам с применением интеллектуальных методов анализа.

Заключение

Таким образом, в результате исследований представлен вариант алгоритмического обеспечения задачи построения структуры web-сайтов. Во взаимосвязи с описанием возможных сценариев компьютерных атак представленный подход способствует идентификации объектов эксплуатации угроз информационной безопасности в структуре web-сайтов, что в целом положительно влияет на их защищенность. Это, в свою очередь, положительно сказывается на стабильности функционирования, в частности web-сайтов государственных органов.

Литература

1. Федеральный закон от 09.02.2009 № 8-ФЗ «Об обеспечении доступа к информации о деятельности государственных органов и органов местного самоуправления».

2. Приказ Минэкономразвития России от 15.11.2022 № 624 «Об утверждении Требований к технологическим, программным и лингвистическим средствам обеспечения пользования официальными сайтами федеральных органов исполнительной власти и подведомственных им организаций».

3. Указ Президента РФ от 05.12.2016 № 646 «Об утверждении Доктрины информационной безопасности Российской Федерации».

4. Синев С.Г., Козачок В.И., Комашинский В.В., Жусов Д.Л. Модель фильтрации потока запросов к web-серверу // Безопасность информационных технологий, № 3, 2007. – С. 75 – 80.

5. Актуальные киберугрозы: IV квартал 2023 года – 2024. URL: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2023-q4/ (дата обращения: 01.12.2024).

6. Результаты мониторинга официальных сайтов федеральных органов исполнительной власти – 2013. URL: http://svobodainfo.org/ru/ node/2527 (дата обращения: 01.12.2024).

7. Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд П. Ривест, Клиффорд Штайн Алгоритмы: построение и анализ – 2-е изд. – М.: Вильямс, 2006. – 1296 с.

8. Низамутдинов М.Ф. Тактика защиты и нападения на web-приложения. – СПб.: БХВ-Петербург, 2005. – 432 с.: ил.

References

1. Federal'nyy zakon ot 09.02.2009 № 8-FZ «Ob obespechenii dostupa k informatsii o deyatel'nosti gosudarstvennykh organov i organov mestnogo samoupravleniya».

2. Prikaz Minekonomrazvitiya Rossii ot 15.11.2022 № 624 «Ob utverzhdenii Trebovaniy k tekhnologicheskim, programmnym i lingvisticheskim sredstvam obespecheniya pol'zovaniya ofitsial'nymi saytami federal'nykh organov ispolnitel'noy vlasti i podvedomstvennykh im organizatsiy».

3. Ukaz Prezidenta RF ot 05.12.2016 № 646 «Ob utverzhdenii Doktriny informatsionnoy bezopasnosti Rossiyskoy Federatsii».

4. Sinev S.G., Kozachok V.I., Komashinskiy V.V., Zhusov D.L. Model' fil'tratsii potoka zaprosov k web-serveru // Bezopasnost' informatsionnykh tekhnologiy, № 3, 2007. – S. 75 – 80.

5. Aktual'nyye kiberugrozy: IV kvartal 2023 goda – 2024. URL: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2023-q4/ (data obrashcheniya: 01.12.2024).

6. Rezul'taty monitoringa ofitsial'nykh saytov federal'nykh organov ispolnitel'noy vlasti – 2013. URL: http://svobodainfo.org/ru/ node/2527 (data obrashcheniya: 01.12.2024).

7. Tomas KH. Kormen, Charl'z I. Leyzerson, Ronal'd P. Rivest, Klifford Shtayn Algoritmy: postroyeniye i analiz – 2-ye izd. – M.: Vil'yams, 2006. – 1296 s. 8. Nizamutdinov M.F. Taktika zashchity i napadeniya na web-prilozheniya. – SPb.: BKHV-Peterburg, 2005. – 432 s.: il.

Жусов Дмитрий Леонидович, кандидат технических наук, доцент, сотрудник федерального государственного казённого военного образовательного учреждения высшего образования «Академия Федеральной службы охраны Российской Федерации». 302020, г. Орёл, ул. Приборостроительная, д. 35. E-mail: d.zhusov@mail.ru.

Макеев Сергей Михайлович, кандидат технических наук, доцент кафедры «Защита информации ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: makeevsm@susu.ru.

Соколов Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой защиты информации ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: sokolovan@susu.ru

Zhusov Dmitry Leonidovich, Candidate of Technical Sciences, Associate Professor, employee of the Academy of the Federal Guard Service of the Russian Federation. 302020, Orel, Priborostroitelnaya St., 35. E-mail: d.zhusov@mail.ru.

Makeev Sergey Mikhailovich, Candidate of Technical Sciences, Associate Professor of Information Security Department, Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: makeevsm@susu.ru.

Sokolov Alexander Nikolayevich, Candidate of Technical Sciences, Associate Professor, Head of Department of Information Security, Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: sokolovan@susu.ru.

УДК 004.056 Вестник УрФО № 4(54) / 2024, с. 37–48

Иванов А. В., Огнев И. А., Селифанов В. В. DOI: 10.14529/secur240405

ВОПРОСЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ АУДИТА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Актуальность исследования: в настоящее время вопросы оценки доверия к субъектам информационного обмена повышают свою актуальность. Такая оценка основывается на оценке доверия к процессам информационной безопасности, принадлежащим субъекту информационного обмена. Среди процессов информационной безопасности, подвергающихся оценке доверия, присутствует процесс аудита информационной безопасности. Оценка доверия к процессу аудита информационной безопасности заключается в вычислении трех показателей – показателя структурной целостности, показателя эффективности и показателя зрелости. Целью работы является разработка методики оценки эффективности процесса аудита информационной безопасности на основе анализа функционирования процесса аудита. Используемые методы и технологии: в работе использованы методы математической статистики, имитационного моделирования. Результат: в результате исследования была сформирована методика оценка эффективности процесса аудита информационной безопасности на базе следующих показателей: коэффициента выявления свидетельств аудита, коэффициента соответствия запланированных ресурсов аудита в программе аудита фактически использованным, коэффициента достаточности времени для проведения аудита. Данные показатели были выбраны в силу того, что из всех возможных составляющих процесса аудита только они зависимы от самого процесса аудита, в то время как часть аспектов регулируется за рамками аудита. Практическая значимость: такой подход позволяет провести оценку эффективности процесса аудита информационной безопасности, основываясь на эффективности использования выделенных временных и человеческих ресурсов для достижения целей аудита с наибольшей вероятностью. Полученный результат является частью методики оценки доверия к процессам аудита информационной безопасности и служит в данной методике характеристикой работы процесса аудита во времени.

Ключевые слова: *эффективность, эффективность процесса, эффективность аудита, аудит, доверие, оценка доверия, доверенное взаимодействие, информационная безопасность*

ISSUES OF ASSESSING THE EFFECTIVENESS OF INFORMATION SECURITY AUDIT

Relevance of the study: currently, the issues of assessing trust in the subjects of information exchange are increasing their relevance. Such an assessment is based on the assessment of trust in the information security processes belonging to the subject of information exchange. Among the information security processes subject to trust assessment, there is the information security audit process. Assessing trust in the information security audit process consists in calculating three indicators - the structural integrity indicator, the efficiency indicator and the maturity indicator. The purpose of the work is to develop a methodology for assessing the effectiveness of the information security audit process based on the analysis of the audit process. Methods and technologies used: the work uses the methods of mathematical statistics, simulation modeling. Result: as a result of the study, a methodology for assessing the effectiveness of the information security audit process was formed based on the following indicators: the coefficient of detection of audit evidence, the coefficient of compliance of the planned audit resources in the audit program with those actually used, the coefficient of sufficiency of time for the audit. These indicators were chosen due to the fact that of all the possible components of the audit process, only they depend on the audit process itself, while some aspects are regulated outside the audit. Practical significance: this approach allows for an assessment of the effectiveness of the information security audit process based on the effectiveness of the use of allocated time and human resources to achieve the audit objectives with the greatest probability. The result obtained is part of the methodology for assessing trust in information security audit processes and serves in this methodology as a characteristic of the audit process over time.

Keywords: *efficiency, process efficiency, audit efficiency, audit, trust, trust assessment, trusted interaction, information security*

Введение

Аудит информационной безопасности представляет собой процесс сбора свидетельств деятельности определенного процесса или группы процессов и проверку этих свидетельств на соответствие требованиям, выдвигаемых к исследуемым процессам или группам процессов [1]. В рамках области информационной безопасности объектом процесса аудита зачастую являются системы защиты информации, процессы управления информационной безопасностью, а также иные процессы или их части, связанные с деятельностью по защите информации [2; 3]. То есть процесс аудита информационной безопасности является методом контроля других процессов информационной безопасности [4].

Построение эффективного процесса аудита информационной безопасности напрямую влияет на общий уровень безопасности информации в силу того, что своевременное и качественное выявление отклонений в системах защиты информации и соответствующих процессов от эталонных показателей закладывает начало своевременному устранению таких отклонений, что снижает вероятность реализации угроз безопасности информации из-за различного рода ошибок [5; 6]. Помимо прочего эффективность процесса аудита влияет на использование и планирование использования сил и времени для снижения количества издержек при выполнении аудита и повышения эффективности процесса как такового.

Целью данного исследования является формирование методики оценки эффективности процесса аудита информационной безопасности. В рамках текущей работы были рассмотрены следующие вопросы: формирование объекта и предмета исследования, формальная постановка задачи исследования, формирование показателей эффективности процесса аудита, формирование функции оценки эффективности процесса аудита, демонстрация работы функции оценки эффективности процесса аудита с использованием средств имитационного моделирования.

Описание процесса аудита информационной безопасности

Цель процесса аудита информационной безопасности – выявление несоответствий (нарушений) реализованных мер по защите информации требованиям регуляторов в области информационной безопасности или стандартов [7–10]. В ходе процесса аудита проводится поиск свидетельств аудита, которые касаются реализации или отсутствия реализации мер по защите информации. Затем осуществляется оценка этих свидетельств, направленная на проверку их соответствия установленным требованиям и стандартам относительно каждой отдельной меры, по итогу чего выявляются отклонения от требований или стандартов [11; 12]. Формальное описание процесса аудита информационной безопасности сводится к формированию определенных наборов входных и выходных данных в соответствии с определенными правилами.

В качестве выходных данных процесса аудита информационной безопасности должен появиться набор замечаний Z (обнаружений аудита):

$$Z = \{z_1..z_j\} \quad (1)$$

где j – количество замечаний аудита.

В процессе проведения аудита информационной безопасности экспертная комиссия анализирует ряд свидетельств аудита C, связанных с каждой мерой защиты информации:

$$C = \{c_1..c_M\} \quad (2)$$

где M – количество свидетельств аудита.

Замечания по аудиту (обнаружения аудита) выявляются на основе анализа свидетельств аудита путем выявления нарушений в реализации мер по защите информации $z(c_i)$:

$$z(c_i) = \begin{cases} 0, \exists c_i \wedge c_i \in Y^{mp} \\ 1 \end{cases} \quad (3)$$

где i – порядковый номер требования ЗИ,

c_i – свидетельство аудита, относящееся к i-му требованию ЗИ, $c_i \in C$,

$Y^{тр}$ – набор требований ЗИ:

$$Y^{mp} = \{Y_1^{mp}..Y_i^{mp}\} \quad (4)$$

То есть в случае, если существуют свидетельство аудита, подтверждающее реализацию i-й меры ЗИ и соответствующие требованию ЗИ $Y^{тр}$, то нарушения в реализации меры ЗИ не обнаружено. В иных случаях, если свидетельств аудита не существует или обнаруженные свидетельства аудита не соответствуют эталонным, то считаем, что нарушение в реализации меры ЗИ обнаружено, соответственно появляется замечание, касающееся i-й меры ЗИ, которое должно быть включено в заключение аудита.

Формально создание и наполнение аудиторского заключения (отчета по аудиту) W можно выразить как составной оператор [13]:

$$\psi = U^c \times C \times \Pi \times R \times Y^{mp} \rightarrow W \quad (5)$$

где U^c – множество актуальных (существенных) угроз безопасности информации,

C – множество свидетельств аудита,

Π – множество программ аудита, применимых для текущего объекта аудита,

R – ресурс, выделенный для проведения аудита,

Y^{mp} – проверяемые требования по ЗИ,

При этом аудиторское заключение содержит ряд обнаружений аудита, основанных на выявленных нарушениях в реализации мер ЗИ:

$$W = Z(C) \quad (6)$$

Зависимость выявления замечаний по аудиту от нарушений в реализации мер ЗИ приведено в формуле (3).

Задача исследования

На сегодняшний день не определен единый формат оценки эффективности процесса аудита информационной безопасности [14], а существующие методы оценки эффективно-

сти процесса аудита информационной безопасности представлены следующими группами методов [15]:

- методы, направленные на оценку достаточности обеспечения аудита ресурсами [16–18];
- методы, направленные на оценку зрелости аудита [6; 19; 20];
- методы, направленные на удовлетворенности результатами аудита [16; 21].

Методы первого типа (оценка достаточности обеспечения аудита ресурсами) направлены на оценку достаточности времени, финансов, людей, выделяемых для реализации процессов аудита. В таком виде оценка носит ограниченный характер из-за того, что объект оценки – ресурсы. Как следствие такие методы оценки могут опосредовано свидетельствовать об эффективности непосредственно процесса аудита информационной безопасности.

Методы основанные на оценке зрелости в первую очередь за объект оценки принимают уровень развития процесса аудита и его составляющих: подпроцесса планирования, подпроцесса исполнения, подпроцесса отчетности и ликвидации нарушений. Здесь зачастую идет экспертная оценка с усреднением результатом, что является основным недостатком таких методов. Такая оценка не позволяет сократить время оценки процесса аудита и подвержена влиянию человеческого фактора.

Последний тип методов, направленный на оценку удовлетворенности результатами аудита за основу оценки берет отзывы различных групп людей различных подразделений различного уровня с их мнением по поводу проведенного аудита. Такая оценка является также опосредованной в отношении эффективности аудита и при этом также являются сугубо субъективной оценкой порой без определенных критериев оценки.

Итого существующие методы оценки эффективности процесса аудита, в том числе процесса аудита информационной безопасности за объект оценки принимают узкий набор аспектов аудита, основаны на экспертной оценке с уравновешиваем результатов и не пригодны для автоматизации в целях сильного снижения времени оценки процесса аудита. В данной работе предлагается способ оценки эффективности процесса аудита информационной безопасности на основе оценки критериев, непосредственно зависимых от самого процесса. При этом предлагается аппарат оценки с четко определенными показателями, влияющими на эффективность процесса аудита, с возможностью автоматизации и проведении оценки в срок менее 24 часов.

В рамках настоящего исследования предлагается решение вопроса формирования показателей эффективности процесса аудита, то есть нахождение такого $f(x)$, которое будет демонстрировать эффективность процесса аудита информационной безопасности A:

$$A = f\left(x\right) \tag{7}$$

где $f(x)$ – функция оценки эффективности процесса аудита информационной безопасности.

Количество свидетельств аудита и длительность их оценки напрямую влияют на вероятность достижения цели аудита – выявления всех нарушений требований или стандартов. Недостаточное количество свидетельств аудита, как и малое время их оценки могут являться причинами неполного выявления нарушений требований или стандартов, или же будет достигнута недостаточная достоверность результатов аудита [22; 23].

Итого для достижения цели аудита – формирование заключения по аудиту, содержащего замечания по реализации мер по ЗИ. Формируется данная цель, исходя из ряда параметров:

1. Множества актуальных угроз безопасности информации (УБИ);
2. Множества выявленных свидетельств аудита;
3. Программы аудита;
4. Множества ресурсов, выделенных для аудита;
5. Количества мер по защите информации (ЗИ), подлежащих оценке.

Результаты исследований

Множества, связанные с актуальными угрозами безопасности информации, содержанием программы аудита, количеством проверяемых мер по защиты информации считаются независимыми от непосредственно процесса аудита и членов команды аудита, в силу того, что программа формируется исходя из целей и задач аудита, формируемых совместно с заказчиком или руководством организации, множество актуальных УБИ и мер ЗИ, формируется на этапе создания системы

ЗИ и в процессе аудита происходит лишь верификация УБИ и мер ЗИ, а иногда верификация вообще отсутствует в зависимости от типа и целей аудита.

Далее отдельно остановимся на оставшихся пунктах, которые напрямую зависят от действий, производимых непосредственно на этапе сбора и анализа информации в процессе аудита [15].

Коэффициент нахождения всех свидетельств аудита

Множество выявленных свидетельств аудита является в свою очередь тем самым, что появляется и анализируется в процессе аудита ИБ. От полноты нахождения элементов множества зависит достижение цели аудита. Соответственно коэффициент нахождения элементов в множестве является одним из основных показателей эффективности аудита.

Коэффициенты заполнения вышеуказанных множеств рассчитываются как отношение фактически найденных элементов к максимально возможному количеству:

$$C = \frac{C_m}{M} \tag{8}$$

где C – коэффициент нахождения всех свидетельств аудита,

C_m – количество обнаруженных свидетельств аудита, $c_m \in C,\ m < M$.

Коэффициент соответствия ресурсов запланированным показателям

Программа аудита в составе своем содержит ряд параметров аудита, необходимых для проведения последнего [24]:

- цель аудита;
- границы аудита;
- критерии аудита;
- дорожная карта аудита;
- команда аудита.

В нашем случае интерес представляют последний и предпоследний пункты. Достаточность и избыточность данных параметров на аудит разобрано далее по тексту, здесь остановимся на ином показателе – коэффициент отклонения от программы аудита [25]:

$$O_L = \frac{\left|L_ф - L_n\right|}{L_n} \tag{9}$$

где O_L – коэффициент отклонения по показателю человеческих ресурсов,

$L_ф$ – фактическое количество аудиторов, $0<L_ф \le 2L_п$,

$L_п$ – запланированное количество аудиторов.

$$O_T = \frac{\left|T_ф - T_n\right|}{T_n} \tag{10}$$

где $P(O_T)$ – коэффициент отклонения по показателю временных ресурсов,

$T_ф$ – фактическое количество времени, затраченного на аудит, $0<T_ф \le 2T_п$,

$T_п$ – запланированное количество времени на аудит.

Далее можно определить общую величину коэффициент отклонений от программы аудита:

$$O = O_L * O_T \tag{11}$$

Соответственно показатель отклонения процесса аудита от запланированных объемов сил и времени фиксирует процент отклонений в каждом процессе непосредственно. Применение данного показателя на статистических данных позволит по смыслу перейти на частоту и размер отклонений от программы аудита и коэффициент придерживания запланированным объемам сил и времени в рамках проведения аудита информационной безопасности.

Показатель коэффициент соответствия ресурсов запланированным показателям обеспечения аудита S будет рассчитываться следующим образом:

$$S = 1 - O \tag{12}$$

Коэффициент достаточности временных ресурсов на аудит

Множество ресурсов, выделенных для аудита, можно интерпретировать по-разному, например, как совокупность ресурсов для сбора свидетельств аудита и для реализации анализа свидетельств аудита [13], как совокупность временных ресурсов [23; 26] или совокупность ресурсов человеческих [27; 28]. В данном исследовании аудиторские ресурсы будем интерпретировать как совокупность временных и человеческих ресурсов. В таком виде можно ответить на два вопроса:

1) является ли достаточным выделенное количество времени на проведение аудита?

2) является ли достаточным выделенное количество людей на проведение аудита?

Таким образом данные показатели достаточности ресурсов также являются одними из основных компонент вероятности достижения целей аудита.

Вопрос достаточности времени может быть сформирован как 2 вариации, зависящих от детерминированности времени проведения каждого этапа аудита. В первом случае, если время каждого этапа является детерминированным (известным и точным), то задача выявления времени, достаточного для проведения аудита, достаточно проста:

$$T_{mp} = \sum_{j=1}^{J} T_j \qquad (13)$$

где T_{mp} – время, требуемое для проведения аудита,

T_j – время проведения j-ого этапа аудита,

J – количество этапов аудита.

Такая формула может применяться, если заранее определены наборы времени для проведения каждого этапа аудита. Например, такие наборы времени могут быть прописаны в (типовой) программе аудита, регламентироваться требованиями по ЗИ или стандартами, выведены эмпирическим путем командой по аудиту и т.п. В таком случае физический смысл формулы также достаточно прост – если команда аудита знает количество времени, требуемое для проведения этапов аудита, то, суммируя это время, можно сказать об известности времени, требуемого для проведения полного аудита (что и является целью аудита).

Иногда можно встретить вторую ситуацию – время каждого этапа аудита строго не определено. Тогда примем, что время проведения каждого i-ого этапа аудита Ti является случайной величиной, подчиняющаяся закону нормального распределения. Тогда время, требуемое для аудита, является также случайной величиной с нормальным распределением, при этом математическое ожидание и дисперсия этой величины будет являться суммой математических ожиданий и дисперсий каждого этапа аудита соответственно. Тогда время $T_{тр}$ будет рассчитываться как:

$$T_{mp} = \arg\left(f\left(\frac{T_{ф} - \sum_{j=1}^{J} M\left(T_j\right)}{\sqrt{\sum_{j=1}^{J} \sigma^2\left(T_j\right)}} \right) \middle| f\left(T\right) = P\left(A\right)_{треб} \right) \qquad (14)$$

где *arg()* – аргумент функции *f(T)*,

f(T) – функция вероятности нормального распределения,

$T_{ф}$ – время проведения аудита,

T_j – время проведения этапа аудита,

$M(T_j)$ – математическое ожидание времени проведения j-ого этапа аудита,

$\sigma^2(T_j)$ – дисперсия времени проведения j-ого этапа аудита,

$P(A)_{треб}$ – требуемое значения вероятности достижения целей аудита.

Здесь для расчета необходимо при помощи лица, принимающего решения, определить требуемую вероятность достижения целей аудита и далее путем анализа графика функции нормального распределения определить значение времени, необходимого для достижения целей аудита с заданной вероятностью.

Формулу (13) можно использоваться в 2 целях:

• ретроспективный анализ, на основе прошлых аудитов (требуемая вероятность аудита становится фактической рассчитанной на основе уже имеющихся данных);

• прогнозирование (лицо, принимающее решение, может задать желаемую вероятность и тогда на базе этого можно провести анализ достаточности времени).

Коэффициент достаточности времени D_T здесь уже рассчитывается как коэффициент, основанный на разности фактического времени аудита и достаточного:

$$D_T = 1 - \frac{\left|T_{ф} - T_{mp}\right|}{T_{mp}} \qquad (15)$$

При этом $0 < T\text{ф} \leq 2T_{тр}$.

Здесь знак разницы неважен для расчета эффективности аудита. Если разница принимает положительный знак, то речь идет об избыточности, в ином случае – о недостаточности. При избыточности времени выделенные

ресурсы для проведения аудита задействованы в неполную силу, а также возможна перегрузка основных бизнес-процессов. При недостатке времени будет цель аудита может быть достигнута с недостаточной достоверностью (формула (13).

Человеческий ресурс имеет влияние на время проведения аудита ИБ, часть задач можно делать параллельно, что снижает общее время проведения аудита и наоборот увеличивает, если количество людей ниже. Количество аудиторов имеет влияние на этапе сбора свидетельства аудита и выявления обнаружений аудита, так как данные 2 этапа являются основными в ходе процесса аудита ИБ. Остальные этапы по большей части независимы от количества аудиторов, поэтому здесь пренебрежем зависимостью времени аудита от человеческих ресурсов в силу значительно малой величины.

Вербально влияние может быть описано следующим образом: если можно выполнять часть независимых этапов (подэтапов) аудита параллельно, то можно сгруппировать время, необходимое для выполнения этих этапов, по принципу максимального значения. Математически мы переходим от суммирования времени этапов аудита к поиску максимального значения:

$$T_{j1-jk} = \begin{cases} T_{j1} + T_{j2} + \ldots + T_{jk}, \text{у} V \vee L < 2 \\ \max\left(T_{j1}, T_{j2}, \ldots, T_{jk}\right), \exists V \wedge L \geq 2 \end{cases} \quad (16)$$

где $T_{j1\text{-}jk}$ – общее время, требуемое для проведения этапов аудита j_1-j_k,

T_{jk} – время, требуемое для проведения j_k-ого этапа аудита,

V – факт возможности распараллеливания проведения этапов аудита,

L – количество аудиторов.

Таким образом человеческий ресурс влияет на временной ресурс аудита, необходимый для достижения целей аудита с требуемой вероятностью.

Коэффициент эффективности аудита

Учитывая все вышесказанное коэффициент эффективности аудита А будет находиться как произведение коэффициент нахождения свидетельств аудита, соответствия ресурсов аудита запланированным показателям и достаточности времени:

$$A = C \times S \times D_T \quad (17)$$

В результате разработанной методики оценки эффективности аудита информационной безопасности появляется возможность получить следующие результаты:

– статистические характеристики оцениваемых показателей (текущее, минимальное, максимальное, среднее значение);

– графики текущих и средних значений оцениваемых показателей;

– гистограммы и статистические функции распределения значений оцениваемых показателей, которые позволят провести опосредованную оценку следующих вопросов:

• управления временными и человеческими ресурсами;

• формирования ретроспективной оценки процесса аудита в течение длительного времени;

• скорости реакции процесса аудита на изменения в процессе аудита.

Имитационное моделирование

На основе сформированной эталонной модели процесса аудита информационной безопасности [5] можно сформировать модель (рис. 1) для расчета показателей эффективности аудита. В качестве среды имитационного моделирования используется программный комплекс AnyLogic, который предоставляет возможность имитации и анализа процессов информационной безопасности.

Здесь реализованы все основные процедуры процесса аудита информационной безопасности, отвечающие за сбор и анализ свидетельств аудита. Далее сформируем ряд экспериментов, касающихся реализации оценки эффективности. Примем следующие показатели за постоянные:

1) планируемое количество аудиторов $L_{\text{п}} = 5$ (чел.);

2) планируемое количество времени $T_{\text{п}} = 19$ (дней);

3) требуемое количество времени $T_{mp} = 20$ (дней);

4) максимальное количество свидетельств аудита $M = 30$ (шт.).

Далее на сымитированных статистических данных проведем оценку процесса аудита информационной безопасности на базе нескольких прогонов системы:

1) фактическое количество свидетельств аудита $C_m = [20..30]$;

2) фактическое количество аудиторов $L_{\text{ф}} = [3..7]$;

3) фактическое количество времени $T_{\text{ф}} = [14..28]$.

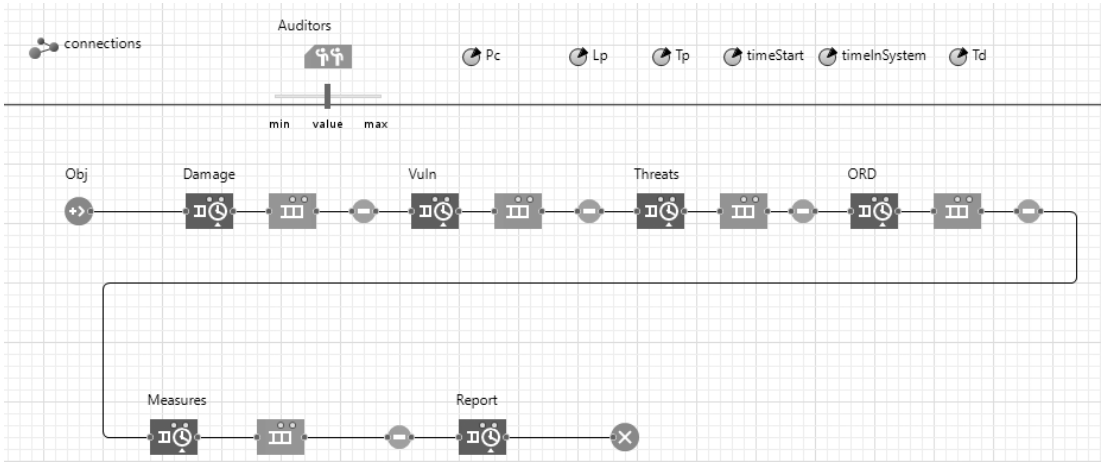


Рис. 1. Имитационная модель процесса аудита

Результаты прогона модели представлены на рисунке 2.

Анализ представленных графиков зависимости эффективности процесса аудита от таких ключевых показателей, как коэффициент нахождения всех свидетельств аудита, коэффициент соответствия ресурсов аудита запланированному количеству в программе аудита и коэффициент достаточности времени, выделенного для аудита, свидетельствует о критической важности поддержания баланса этих факторов. Отклонения любого из указанных показателей в любую сторону ведут к снижению общей эффективности процесса аудита. Например, недостаточное количество обнаруженных свидетельств аудита или избыточное отклонение от запланированной программы аудита могут приводить к пропуску важных нарушений или к нецелевому расходу ресурсов. С другой стороны, как недостаточное, так и избыточное время, выделенное на аудиторские процедуры, приводит либо к поверхностному анализу, либо к растрате ресурсов, не влияя на улучшение качества процесса аудита информационной безопасности. Таким образом, поддержание баланса между этими показателями является решающим для достижения оптимальной эффективности аудита информационной безопасности.

Заключение

В результате проведенного исследования были рассмотрены вопросы проведения оценки эффективности процесса аудита информационной безопасности. За показатели эффективности были приняты следующие аспекты аудита – коэффициент нахождения всех свидетельств аудита, коэффициент соответствия использованных временных и человеческих ресурсов запланированным в программе аудита, коэффициент достаточности времени, выделенного для проведения аудита информационной безопасности.

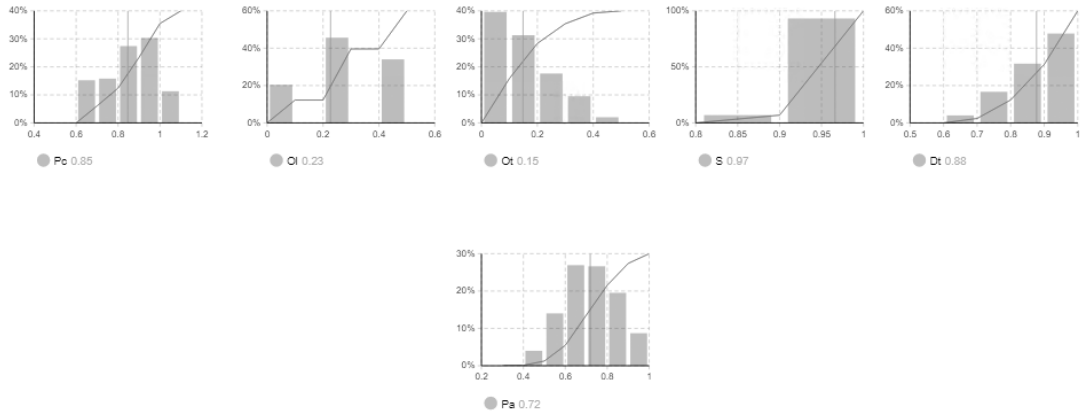


Рис. 2. Статистика измерений показателей эффективности аудита

Сформированный механизм оценки эффективности аудита информационной безопасности призван оценивать эффективность использования выделенных временных и человеческих ресурсов для достижения целей аудита. Показатель эффективности процесса аудита будет использоваться в целях проведения оценки доверия к процессу аудита информационной безопасности как мера работоспособности аудита в рамках реализации мероприятий по обеспечению информационной безопасности.

Данная работа выполнена при финансовой поддержке Фонда поддержки проектов Национальной технологической инициативы (НТИ) в рамках реализации Программы Центра компетенций НТИ «Технологии доверенного взаимодействия» (договор от «14» декабря 2021 г. № 70-2021-00246).

Литература

1. Information security audit for a manufacturing company / S.V. Shirokova [et al.] // Information and control systems. – 2023. – Vol. 122. – № 1. – P. 41-50.

2. Каширская Л.В. Объекты аудита информационной безопасности и направления их проверки / Л.В. Каширская, Ю.А. Зурнаджьянц // Аудитор. – 2022. – Vol. 8. – № 1. – P. 21-31.

3. Senkiv D.A. Audit as a Means of Ensuring Information Security of Web Applications and Used Computer Systems / D.A. Senkiv // American Scientific Journal. – 2020. – Vol. 40. – № 2. – P. 54-57.

4. Santi R. Information system security audit using ISO/IEC 27002:2013 at university of XXX / R. Santi, A.I. Alfresi, B. Octariana // Jurnal Teknik Informatika (Jutif). – 2023. – Vol. 4. – № 4. – P. 733-750.

5. Огнев И.А. Вопросы математической интерпретации процесса аудита информационной безопасности с применением сетей Петри / И.А. Огнев // Доклады Томского Государственного Университета Систем Управления И Радиоэлектроники. – 2024. – Vol. 27. – № 2. – P. 15-20.

6. Effectiveness of cybersecurity audit / S. Slapničar [et al.] // International Journal of Accounting Information Systems. – 2022. – Vol. 44. – P. 100548.

7. Макаренко С.И. Аудит безопасности критической информационной инфраструктуры / С.И. Макаренко. – СПб: Издательство «Наукоемкие технологии», 2023. – 122 p.

8. Денисенко В.В. Аудит Информационной Безопасности Организаций: Методы И Преимущества / В.В. Денисенко, А.М. Гончаров, И.П. Маслов // Наукосфера. – 2023. – № 11-2. – P. 135-140.

9. Ситская А.В. Вопросы аудита информационной безопасности / А.В. Ситская, В.В. Селифанов, П.А. Звягинцева // Безопасность цифровых технологий. – 2023. – Vol. 110. – № 3. – P. 67-82.

10. Чекулаева Е.Н. Методика аудита информационной безопасности предприятия с использованием причинно-следственной диаграммы / Е.Н. Чекулаева, Е.С. Кубашева // Вестник Поволжского государственного технологического университета. Серия: Радиотехнические и Инфокоммуникационные системы. – 2020. – Vol. 45. – № 1. – P. 58-68.

11. Héroux S. The internal audit function in information technology governance: A holistic perspective / S. Héroux, A. Fortin // Journal of Information Systems. – 2013. – Vol. 27. – № 1. – P. 189-217.

12. IT governance and IT controls: Analysis from an internal auditing perspective / T.-H. Wu [et al.] // International Journal of Accounting Information Systems. – 2024. – Vol. 52. – P. 100663.

13. Воеводин В.А. Концептуальная модель объекта аудита информационной безопасности / В.А. Воеводин // Computational nanotechnology. – 2019. – № 3. – P. 92-95.

14. Kotb A. Mapping of Internal Audit Research: A Post-Enron Structured Literature Review / A. Kotb, H. Elbardan, H. Halabi // Accounting Auditing & Accountability Journal. – 2020. – Vol. 33. – № 8. – P. 1969-1996.

15. Turetken O. Internal audit effectiveness: operationalization and influencing factors / O. Turetken, S. Jethefer, B. Ozkan // Managerial Auditing Journal. – 2020. – Vol. 35. – № 2. – P. 238-271.

16. Constructing internal audit quality evaluation index: evidence from listed companies in Jiangsu province, China / R. Kai [et al.] // Heliyon. – 2022. – Vol. 8. – № 9. – P. e10598.

17. Examining the critical factors of internal audit effectiveness from internal auditors' perspective: Moderating role of extrinsic rewards / H. Alqudah [et al.] // Heliyon. – 2023. – Vol. 9. – № 10. – P. e20497.

18. Ефремов А.В. Анализ существующих методик оценки средств аудита информационной безопасности / Ефремов А.В., Панамарев Г.Е. // ВЕСТНИК ВОЕННОГО ИННОВАЦИОННОГО ТЕХНОПОЛИСА "ЭРА." – 2021. – Vol. 2. – № 4. – P. 38-45.

19. Gaosong Q. Measurement of Internal Audit Effectiveness: Construction of Index System and Empirical Analysis / Q. Gaosong, Y. Leping // Microprocessors and Microsystems. – 2021. – P. 104046.

20. Roussy M. Internal audit: from effectiveness to organizational significance / M. Roussy, O. Barbe, S. Raimbault // Managerial Auditing Journal. – 2020. – Vol. 35. – № 2. – P. 322-342.

21. Сафохина Е.А. Эффективность внутреннего аудита как элемент обеспечения экономической безопасности хозяйствующего субъекта / Е.А. Сафохина // Вестник экономической безопасности. – 2022. – № 1. – P. 301-306.

22. Hutchinson B. Audit masquerade: How audits provide comfort rather than treatment for serious safety problems / B. Hutchinson, S. Dekker, A. Rae // Safety Science. – 2024. – Vol. 169. – P. 106348.

23. Макаренко С.И. Критерии и показатели оценки качества тестирования на проникновение / С.И. Макаренко // Вопросы кибербезопасности. – 2021. – Vol. 43. – № 3. – P. 43-57.

24. ISO 19011:2018 - Guidelines for auditing management systems.

25. Бусуёк Н.А. Аудит эффективности в системе внешнего государственного финансового контроля (аудита) / Н.А. Бусуёк, Л.М. Макарова // Вестник Московского финансово-юридического университета. – 2022. – № 3. – P. 140-145.

26. Calabrese K. The effects of time pressure on audit fees / K. Calabrese // Advances in Accounting. – 2023. – Vol. 63. – P. 100663.

27. Воеводин В.А. Определение весомости аудиторских свидетельств методом бальных оценок при аудите информационной безопасности / В.А. Воеводин, М.С. Маркина, П.В. Маркин // Computational nanotechnology. – 2020. – № 1. – P. 57-62.

28. Opening the black box of human resource allocations in audit firms: The assignment of audit partners to audit engagements / B. Wu [et al.] // The British Accounting Review. – 2024. – Vol. 56. – № 2. – P. 101231.

References

1. Shirokova S.V., Rostova O.V., Bolsunovskaya M.V., Dmitrieva L.A., Almataev T.O. Information security audit for a manufacturing company. Information and control systems, 2023, vol. 122, no. 1, pp. 41-50.

2. Kashirskaya L.V., Zurnadzh'yants Yu.A. Ob"ekty audita informatsionnoi bezopasnosti i napravleniya ikh proverki [Objects of information security audit and directions of their verification]. Auditor, 2022, vol. 8, no. 1, pp. 21-31. (In Russ.)

3. Senkiv D.A. Audit as a Means of Ensuring Information Security of Web Applications and Used Computer Systems. American Scientific Journal, 2020, vol. 40, no. 2, pp. 54-57.

4. Santi R., Alfresi A.I., Octariana B. Information system security audit using ISO/IEC 27002:2013 at university of XXX. Jurnal Teknik Informatika (Jutif), 2023, vol. 4, no. 4, pp. 733-750.

5. Ognev I.A. Voprosy matematicheskoi interpretatsii protsessa audita informatsionnoi bezopasnosti s primeneniem setei Petri [Issues of mathematical interpretation of the information security audit process using Petri nets]. Reports of Tomsk State University of Control Systems and Radioelectronics, 2024, vol. 27, no. 2, pp. 15-20. (In Russ.)

6. Slapničar S., Vuko T., Čular M., Drašček M. Effectiveness of cybersecurity audit. International Journal of Accounting Information Systems, 2022, vol. 44, pp. 100548.

7. Makarenko S.I. Audit bezopasnosti kriticheskoi informatsionnoi infrastruktury [Security audit of critical information infrastructure]. SPb: Izdatel'stvo «Naukoemkie tekhnologii», 2023. — 122 p. (In Russ.)

8. Denisenko V.V., Goncharov A.M., Maslov I.P. Audit Informatsionnoi Bezopasnosti Organizatsii: Metody I Preimushchestva [Information Security Audit of Organizations: Methods and Advantages]. Naukosphere, 2023, no. 11-2, pp. 135-140. (In Russ.)

9. Sitskaya A.V., Selifanov V.V., Zvyagintseva P.A. Voprosy audita informatsionnoi bezopasnosti [Information security audit issues]. Digital Technology Security, 2023, vol. 110, no. 3, pp. 67-82. (In Russ.)

10. Chekulaeva E.N., Kubasheva E.S. Metodika audita informatsionnoi bezopasnosti predpriyatiya s ispol'zovaniem prichinno-sledstvennoi diagrammy [Methodology of enterprise information security audit using a cause-and-effect diagram]. Bulletin of the Volga State Technological University. Series: Radio Engineering and Infocommunication Systems, 2020, vol. 45, no. 1, pp. 58-68. (In Russ.)

11. Héroux S., Fortin A. The internal audit function in information technology governance: A holistic perspective. Journal of Information Systems, 2013, vol. 27, no. 1, pp. 189-217.

12. Wu T.-H., Huang S.Y., Chiu A.-A., Yen D.C. IT governance and IT controls: Analysis from an internal auditing perspective. International Journal of Accounting Information Systems, 2024, vol. 52, pp. 100663.

13. Voevodin V.A. Kontseptual'naya model' ob"ekta audita informatsionnoi bezopasnosti [Conceptual model of the object of information security audit]. Computational nanotechnology, 2019, no. 3, pp. 92-95. (In Russ.)

14. Kotb A., Elbardan H., Halabi H. Mapping of Internal Audit Research: A Post-Enron Structured Literature Review. Accounting Auditing & Accountability Journal, 2020, vol. 33, no. 8, pp. 1969-1996.

15. Turetken O., Jethefer S., Ozkan B. Internal audit effectiveness: operationalization and influencing factors. Managerial Auditing Journal, 2020, vol. 35, no. 2, pp. 238-271.

16. Kai R., Yusheng K., Ntarmah A.H., Ti C. Constructing internal audit quality evaluation index: evidence from listed companies in Jiangsu province, China. Heliyon, 2022, vol. 8, no. 9, pp. e10598.

17. Alqudah H., Amran N.A., Hassan H., Lutfi A., Alessa N., alrawad M., Almaiah M.A. Examining the critical factors of internal audit effectiveness from internal auditors' perspective: Moderating role of extrinsic rewards. Heliyon, 2023, vol. 9, no. 10, pp. e20497.

18. Efremov A.V., Panamarev G.E. Analiz sushchestvuyushchikh metodik otsenki sredstv audita informatsionnoi bezopasnosti [Analysis of existing methods for assessing information security audit tools]. BULLETIN OF THE MILITARY INNOVATIVE TECHNOPOLIS "ERA.", 2021, vol. 2, no. 4, pp. 38-45. (In Russ.)

19. Gaosong Q., Leping Y. Measurement of Internal Audit Effectiveness: Construction of Index System and Empirical Analysis. Microprocessors and Microsystems, 2021, pp. 104046.

20. Roussy M., Barbe O., Raimbault S. Internal audit: from effectiveness to organizational significance. Managerial Auditing Journal, 2020, vol. 35, no. 2, pp. 322-342.

21. Safokhina E.A. Effektivnost' vnutrennego audita kak element obespecheniya ekonomicheskoi bezopasnosti khozyaistvuyushchego sub"ekta [Effectiveness of internal audit as an element of ensuring economic security of an economic entity]. Bulletin of Economic Security, 2022, no. 1, pp. 301-306. (In Russ.)

22. Hutchinson B., Dekker S., Rae A. Audit masquerade: How audits provide comfort rather than treatment for serious safety problems. Safety Science, 2024, vol. 169, pp. 106348.

23. Makarenko S.I. Kriterii i pokazateli otsenki kachestva testirovaniya na proniknovenie [Criteria and indicators for assessing the quality of penetration testing]. Cybersecurity Issues, 2021, vol. 43, no. 3, pp. 43-57. (In Russ.)

24. ISO 19011:2018 - Guidelines for auditing management systems.

25. Busuek N.A., Makarova L.M. Audit effektivnosti v sisteme vneshnego gosudarstvennogo finansovogo kontrolya (audita) [Performance audit in the system of external state financial control (audit)]. Bulletin of the Moscow Financial and Law University, 2022, no. 3, pp. 140-145. (In Russ.)

26. Calabrese K. The effects of time pressure on audit fees. Advances in Accounting, 2023, vol. 63, pp. 100663.

27. Voevodin V.A., Markina M.S., Markin P.V. Opredelenie vesomosti auditorskikh svidetel'stv metodom bal'nykh otsenok pri audite informatsionnoi bezopasnosti [Determining the weight of audit evidence using the scoring method in information security audit]. Computational nanotechnology, 2020, no. 1, pp. 57-62. (In Russ.)

28. Wu B., Wu Y., Zhang M., Li J. Opening the black box of human resource allocations in audit firms: The assignment of audit partners to audit engagements. The British Accounting Review, 2024, vol. 56, no. 2, pp. 101231.

Иванов Андрей Валерьевич, кандидат технических наук, доцент, заведующий кафедрой защиты информации, Новосибирский государственный технический университет. Область научных интересов: доверенное взаимодействие, защита информации. 630073, г. Новосибирск, проспект К. Маркса, 20. E-mail: andrej.ivanov@corp.nstu.ru.

Огнев Игорь Александрович, старший преподаватель кафедры защиты информации, Новосибирский государственный технический университет. Область научных интересов: оценка доверия, аудит информационной безопасности. 630073, г. Новосибирск, проспект К. Маркса, 20. E-mail: i.ognev@corp.nstu.ru.

Селифанов Валентин Валерьевич, старший преподаватель кафедры защиты информации, Новосибирский государственный технический университет. Область научных интересов: оценка доверия, управление информационной безопасностью. 630073, г. Новосибирск, проспект К. Маркса, 20. E-mail: selifanov@corp.nstu.ru.

Ivanov Andrey Valerievich, Candidate of Technical Sciences, Associate Professor, Head of the Information Security Department, Novosibirsk State Technical University. Research interests: trusted interaction, information security. 630073, Novosibirsk, avenue K. Marksa, 20. E-mail: andrej.ivanov@corp.nstu.ru.

Ognev Igor Aleksandrovich, Senior Lecturer, Information Security Department, Novosibirsk State Technical University. Research interests: trust assessment, information security audit. 630073, Novosibirsk, avenue K. Marksa, 20. E-mail: i.ognev@corp.nstu.ru.

Selifanov Valentin Valerievich, Senior Lecturer, Information Security Department, Novosibirsk State Technical University. Research interests: trust assessment, information security management. 630073, Novosibirsk, avenue K. Marksa, 20. E-mail: selifanov@corp.nstu.ru.

УДК 004.056 Вестник УрФО № 4(54) / 2024, с. 49–60

Плетенкова А. Д., Соколов А. Н. DOI: 10.14529/secur240406

ПРИМЕНЕНИЕ ДВУХЭТАПНОГО МЕТОДА КЛАСТЕРИЗАЦИИ НА ОСНОВЕ САМООРГАНИЗУЮЩЕЙСЯ КАРТЫ КОХОНЕНА ДЛЯ ОБНАРУЖЕНИЯ АНОМАЛИЙ В СИНТЕТИЧЕСКИХ НАБОРАХ ДАННЫХ[1]

В статье представлен двухэтапный метод кластеризации, основанный на применении самоорганизующейся карты Кохонена с целью выявления аномалий в синтетических наборах данных. Этот подход позволяет более точно идентифицировать аномальные объекты по сравнению с одноэтапными методами кластеризации. Представлены результаты экспериментов, проведенных на синтетических наборах данных, которые подтверждают эффективность предложенного метода.

На первом этапе исследования формируется синтетический набор данных, содержащий два или три признака. Далее данные подвергаются обработке с использованием самоорганизующейся карты Кохонена, что позволяет выделить основные кластеры и определить границы между ними.

На втором этапе обнаружения аномалий применяются алгоритмы кластеризации, основанные на анализе пространства признаков и сравнении расстояний между объектами. Аномальные объекты, находящиеся в значительном удалении от основных кластеров, классифицируются как аномалии. Проведенный анализ работы алгоритмов кластеризации на данном этапе позволил выявить, что алгоритмы DBSCAN и Isolation Forest проявили себя наиболее эффективными в обнаружении выбросов по сравнению с алгоритмом OPTICS.

Ключевые слова: *обнаружение аномалий, кластеризация данных, самоорганизующаяся карта Кохонена, синтетический набор данных*

[1] Исследование поддержано грантом Российского научного фонда (проект No 22-71-10095).

APPLICATION OF A TWO-STAGE CLUSTERING METHOD BASED ON SELF-ORGANISING KOHONEN MAP FOR ANOMALY DETECTION IN SYNTHETIC DATASETS

The paper presents a two-stage clustering method based on the application of a self-organizing Kohonen map to identify anomalies in synthetic datasets. This approach allows for more accurate identification of anomalous objects compared to one-stage clustering methods. Experimental results on synthetic datasets are presented, confirming the effectiveness of the proposed method.

In the first stage of the study, a synthetic dataset containing two or three features is generated. The data is then processed using a self-organizing Kohonen map to identify the main clusters and determine the boundaries between them.

In the second stage of anomaly detection, clustering algorithms based on feature space analysis and distance comparison between objects are applied. Anomalous objects that are far from the main clusters are classified as anomalies. Analysis of the performance of the clustering algorithms at this stage showed that the DBSCAN and Isolation Forest algorithms were the most effective in detecting outliers compared to the OPTICS algorithm.

Keywords: *anomaly detection, data clustering, a self-organizing Kohonen map, synthetic dataset*

Введение

Техническое развитие промышленных средств автоматизации и увеличение уровня интеграции промышленных и корпоративных сетей приводит к увеличению рисков проведения успешных кибератак. Реализация таких кибератак может подразумевать получение доступа к управлению важными промышленными установками, что влечёт за собой риск остановки производства или создания аварийной ситуации.

Обеспечение информационной безопасности автоматизированных систем управления технологическими процессами (АСУ ТП) требует своевременного обнаружения кибератак как известного, так и неизвестного типа. Эти кибератаки можно рассматривать как аномалии в динамических процессах, регистрируемых при работе АСУ ТП [1].

Для решения задачи поиска кластеров произвольной формы в условиях сильной зашумленности в данных обычно применяют простые и быстрые алгоритмы, основанные на методах обучения без учителя, такие, например, как метод K-средних. Однако они неэффективны в условиях сильной зашумленности и требуют дополнительной информации, включая количество кластеров [2].

Более сложные алгоритмы иерархической кластеризации, такие, например, как OPTICS (Ordering Points To Identify the Clustering Structure), определяют кластеры произвольной формы, сохраняя информацию о всей структуре кластеров, однако требуют значительных вычислительных ресурсов при работе с высокоразмерными данными из-за сложности алгоритма [3].

Среди классических алгоритмов кластеризации наиболее перспективным алгоритмом, способным выделять кластеры произвольной формы, является DBSCAN (Density-Based Spatial Clustering of Applications with Noise) – алгоритм пространственной кластеризации на основе плотности данных. Однако следует отметить, что при увеличении количества признаков эффективность DBSCAN

может снижаться. При этом по сравнению с другими алгоритмами кластеризации этот алгоритм обладает сравнительно небольшой вычислительной сложностью [4].

При анализе данных, характеризующихся высокой размерностью, возникает необходимость в таком алгоритме кластеризации, как Isolation Forest. Это алгоритм поиска аномалий, который может использоваться и для кластеризации. В контексте высокой размерности данных Isolation Forest проявляет себя эффективно, поскольку он строит деревья решений быстро и анализирует отдельные объекты независимо от остальных, что позволяет ему хорошо справляться с данными высокой размерности, где другие методы могут столкнуться с проблемами из-за т.н. «проклятия размерности» [5].

Тем не менее, приведенные методы одноэтапного обнаружения аномалий, основанные на использовании статических пороговых значений и подверженные воздействию шума, могут оказаться ограниченными в своей гибкости, устойчивости и способности выявлять сложные аномалии в данных. С целью совершенствования этих методов, расширения их функциональных возможностей и адаптации под задачи эффективного обнаружения и предотвращения кибератак предлагается применять двухэтапный метод кластеризации, основанный на применении самоорганизующейся карты Кохонена.

Для получения точных и устойчивых решений при решении сложных задач в области обнаружения аномалий с применением двухэтапных методов кластеризации включаются такие этапы, как предварительная обработка данных, настройка параметров используемых алгоритмов, комбинирование различных моделей и учет контекста [6].

Обнаруженные аномалии в данных могут свидетельствовать о потенциальных угрозах и атаках и проявляться в различных формах, например, в виде необычных паттернов доступа к данным, отклонений от типичного поведения пользователей или изменения в структуре данных. Процесс кластеризации позволяет группировать данные в соответствии с их сходством по характеристикам, что в свою очередь способствует выявлению аномалий, отличающихся от общей тенденции данных [7].

Целью применения двухэтапного метода кластеризации, основанного на самоорганизующейся карте Кохонена, является формирование кластеров, наблюдаемых данных на первом этапе и выявление возможных угроз безопасности информации на втором этапе. Исследование проводилось с использованием онлайн-платформы Colaboratory от Google на языке программирования Python.

На первом этапе метода происходит построение самоорганизующейся карты Кохонена, которая представляет собой инструмент для визуализации многомерного пространства параметров. В процессе обучения карта формирует топологическую структуру, отражающую взаимосвязи между различными параметрами данных [8].

На втором этапе анализа результатов кластеризации, полученных от самоорганизующейся карты Кохонена, применяются алгоритмы кластерного анализа, такие как DBSCAN, OPTICS и Isolation Forest. В результате обработки данных этими алгоритмами отмечаются однозначно выделенные аномалии, что позволяет эффективно обнаруживать аномальные объекты в исследуемом наборе данных.

Описание набора данных

В работе проведен анализ синтетических двухмерных и трёхмерных наборов данных с бинарной классификацией. Предполагалось, что в данных присутствует заданное количество выбросов, связанное с параметром «contamination» (который составляет 10% от общего объема данных).

Рассмотрим синтетический массив двухмерных данных, соответствующий наличию двух признаков в потоке данных. В массиве содержится 110 точек, где 100 точек классифицируются как «нормальные» (inliers), а 10 точек представляют собой данные, отличающиеся от общей структуры, и считаются «выбросами» (outliers, рис. 1).

```
Первые несколько inliers:
[[ 0.1, 0.9],
 [-0.7, 0.2],
 [ 0.4, -0.8],
 ... ]
```

```
Первые несколько outliers:
[[ -5.2,  3.1],
 [  0.5, -7.3],
 [ -8.0,  0.0],
 ... ]
```

Рис. 1. Представление массива синтетического набора данных

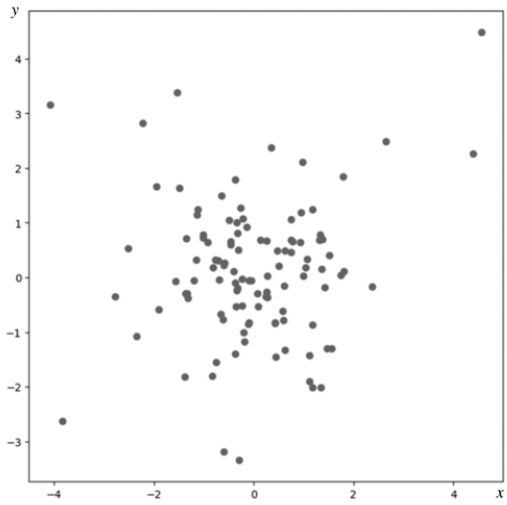


Рис. 2. Диаграмма рассеяния синтетического набора данных

В ходе выполнения алгоритма происходит генерация случайных выбросов, которые могут изменяться при каждом запуске программы. Затем формируется диаграмма рассеяния, которая наглядно отображает распределение точек данных по координатам x и y на декартовой плоскости (рис. 2, 8).

Необходимо отметить, что в процессе синтеза данных не задается фиксированное значение радиуса или расстояния, которое определяло бы данные как аномалии. Процесс обнаружения аномалий изменяется динамически в зависимости от структуры данных. При моделировании данных используется распределение внутри двух концентрических окружностей. Такой способ моделирования данных основан на принципе сферической симметрии и позволяет эффективно выявлять аномалии в данных, структура которых соответствует структуре окружающего пространства данных. Равномерное распределение данных по всей поверхности сферы обеспечивает их симметрию относительно центра, что на фоне симметрии упрощает выделение и анализ аномалий [9].

Для визуального представления данных в виде «выбросов» определено пороговое значение ошибки квантования (рис. 3, 9). На графике представлены значения ошибок квантования по оси X, и частота их появления по оси

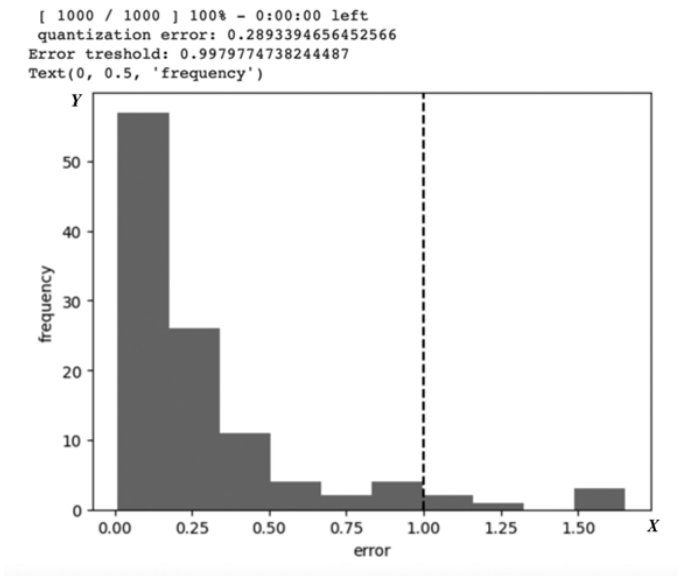


Рис. 3. Гистограмма ошибки квантования

Y. Ось X отражает величину ошибок квантования, т.е. меру расстояния между исходными данными и соответствующими векторами после обучения картой Кохонена. Эти ошибки нормализованы, показывая, насколько далеко точка данных находится от ближайшего нейрона в весовом пространстве. Ось Y отображает частоту точек данных с каждым значением ошибки на оси X. Следовательно, полученная гистограмма визуализирует распределение ошибок в наборе данных. Вертикальная черная пунктирная линия на графике обозначает порог ошибки, превышение которого классифицирует данные как выбросы.

Обычно, для определения порогового значения используются различные статистические методы, такие как, например, методы на основе стандартного отклонения или квантилей [10]. В рассматриваемом случае для вычисления порога выявления аномалий использовано среднее значение ошибок квантования np.mean(quantization_errors) и стандартное отклонение np.std(quantization_errors) этих ошибок. Среднее значение ошибок квантования (np.mean) вычисляется с помощью формулы:

$$\text{mean} = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

где N – общее количество значений, а x_i – значения ошибок квантования.

Стандартное отклонение (np.std) вычисляется с помощью формулы:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2},$$

где μ – среднее значение вероятностного распределения.

Вычисление порога обнаружения аномалий на основе средних значений и стандартных отклонений ошибок квантования осуществлялось по формуле [11]:

$$T = \mu + k \cdot \sigma,$$

где μ = np.mean(quantization_errors) – среднее значение ошибок квантования;

σ = np.std(quantization_errors) – стандартное отклонение ошибок квантования;

k – коэффициент, определяющий порог определения аномалий.

В случае, когда значение ошибки квантования для определенной точки превышает установленный порог T, данная точка классифицируется как аномальная.

Алгоритм двухэтапной кластеризации, основанный на самоорганизующейся карте Кохонена

На первом этапе обработки синтетических данных применена самоорганизующаяся карта Кохонена, которая представляет собой инструмент первоначальной кластеризации данных с целью выявления групп, обладающих схожей структурой данных, отражающих как нормальное, так и аномальное функционирование системы.

Для реализации процесса кластеризации данных построена карта Кохонена размером 8x8 узлов с установленной топологией размещения нейронов. Перед обучением карты произведено нормирование обучающего датасета, а во время обучения использована линейная инициализация весов нейронов и применен пакетный обучающий алгоритм. В контексте карт Кохонена пакетный алгоритм обучения подразумевает использование всех обучающих данных для обновления весов нейронной сети за одну итерацию. Это позволяет более эффективно использовать данные и сделать процесс обучения более стабильным. По завершению обучения карты формируется демонстрационная унифицированная матрица расстояний - представление самоорганизующейся карты Кохонена, которое визуализирует расстояние между нейронами [12].

Исследование начинается с применения модели ко всему набору синтетических данных. Затем каждый входной образец из данных отображается на карту. Для каждого образца вычисляется евклидово расстояние между исходным значением и его аппроксимацией на карте. Темная окраска на карте соответствует увеличению расстояния между нейронами, что свидетельствует о возможном наличии выбросов в данных (рис. 4, 10).

В ходе анализа данных с использованием карты Кохонена были выявлены точки, которые могут быть интерпретированы как выбросы. Проведенное исследование выявило наличие двух кластеров в обученной карте Кохонена. Среди общего числа выбросов, представленных в датасете, было обнаружено 6 точек из 10 заданных (рис. 5, 11).

Эти значения представлены в классификационном отчете и на матрице ошибок, используемой в качестве инструмента визуали-

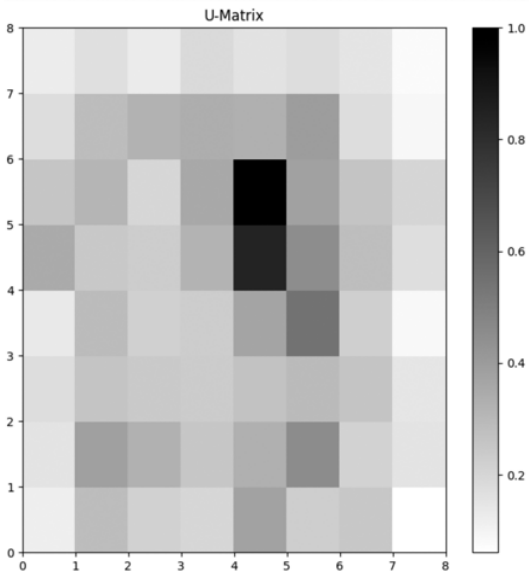


Рис. 4. Унифицированная матрица расстояний (U-matrix) между нейронами

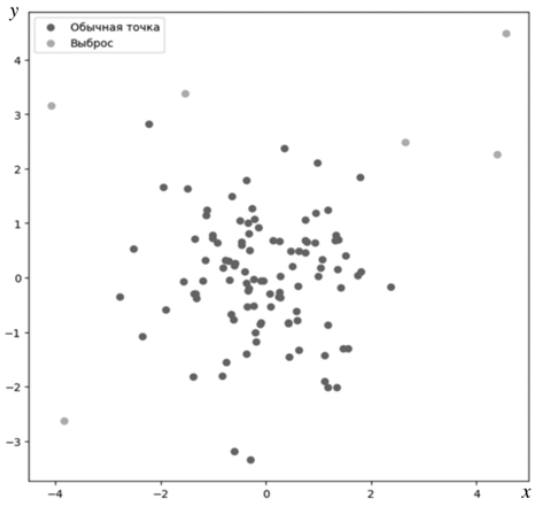
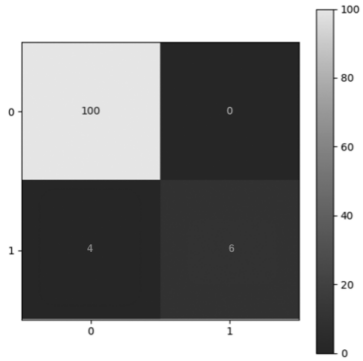


Рис. 5. Диаграмма рассеяния после обработки данных картой Кохонена

зации качества работы классификатора, демонстрирующего, сколько объектов каждого класса было классифицировано правильно или неправильно (рис. 6). Диагональные элементы матрицы отражают количество правильно классифицированных объектов для каждого класса. Ошибки первого и второго рода вычисляются на основе данных из матрицы ошибок, которая содержит информацию о фактических и предсказанных классах.

На втором этапе обработки данных происходит процесс идентификации кластеров, к которым были присвоены данные, описывающие аномальное поведение наблюдаемых процессов на обученной карте. Для выполнения этого этапа исследования применены алгоритмы кластеризации DBSCAN, OPTICS и Isolation Forest (рис. 7, 12).




```
Матрица ошибок:
[[100   0]
 [  4   6]]
Количество False Positive: 0
Количество False Negative: 4
```

а)

	precision	recall	f1-score	support
0.0	0.96	1.00	0.98	100
1.0	1.00	0.60	0.75	10
accuracy			0.96	110
macro avg	0.98	0.80	0.87	110
weighted avg	0.97	0.96	0.96	110

```
Precision: 1.0
Recall: 0.6
F1 Score: 0.7499999999999999
```

б)

Рис. 6. Метрики точности работы алгоритма: матрица ошибок (а); классификационный отчет (б)

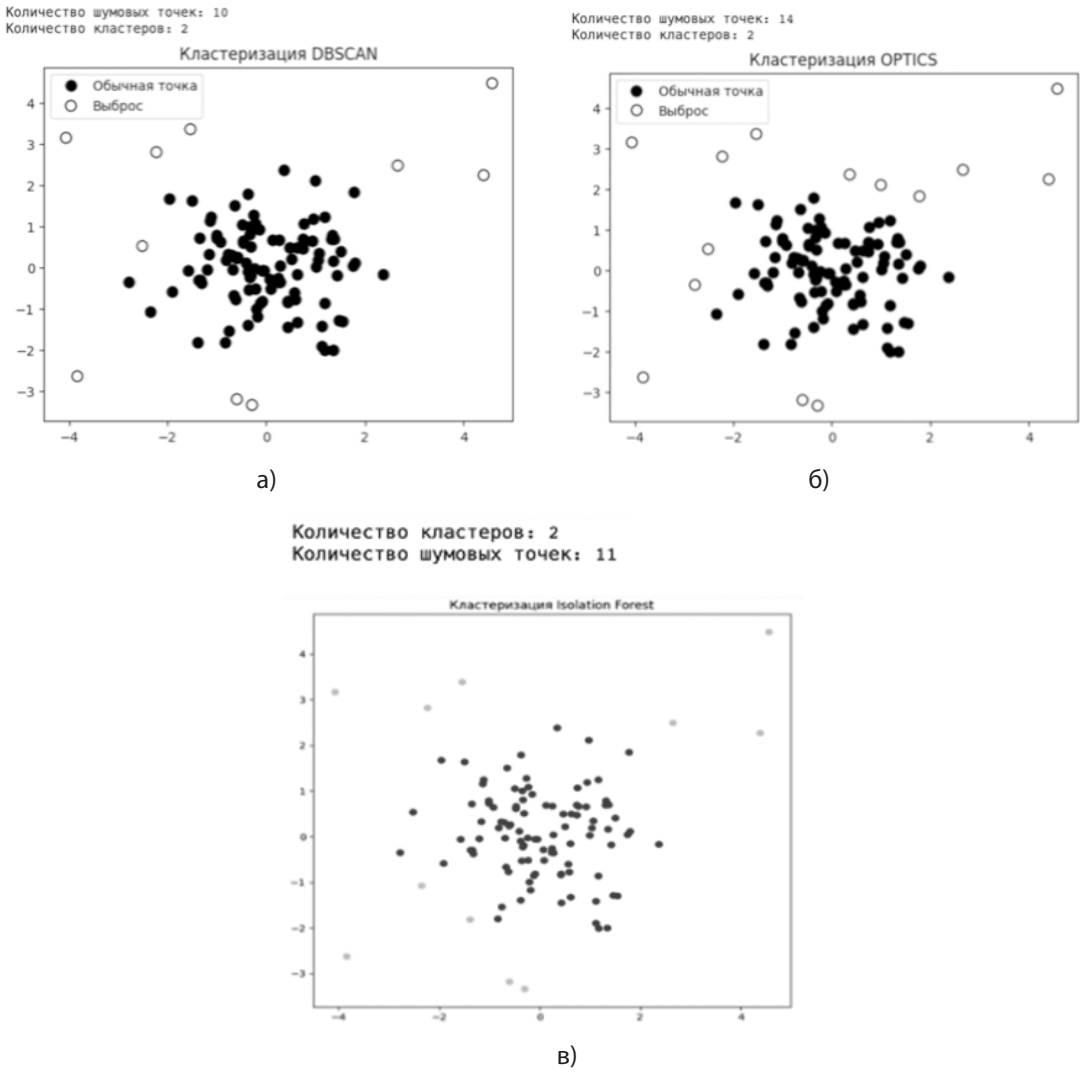


Рис. 7. Диаграммы рассеяния второго этапа кластеризации методами: DBSCAN (а); OPTICS (б) и Isolation Forest (в)

На входе алгоритмов используются векторы весовых коэффициентов, соответствующие каждому из 64 нейронов обученной карты Кохонена.

Рассматривая работу кластеризаторов на втором этапе, можно отметить, что они успешно выявили оставшиеся выбросы. Наиболее результативными при этом оказались методы DBSCAN и Isolation Forest.

Для каждого алгоритма проводилась настройка параметров. Например, настройка работы алгоритма DBSCAN включает в себя два этапа: указание минимального числа соседей, которые ищутся для каждой точки из набора входных данных, и указание радиуса окрестности, в котором ведется соответствующий поиск соседей для каждой точки [2].

Подбор оптимальных параметров был осуществлен с использованием метода перебора. Установленные значения настроечных параметров, такие как радиус окрестности $\varepsilon = 1{,}0$ и минимальное число соседей $m = 8$, использовались для выполнения алгоритма DBSCAN после предварительного обучения картой Кохонена. Результатом работы алгоритма было разделение исходных данных на кластеры, где оранжевые точки были определены как выбросы, а голубая область представляла собой единственный обнаруженный кластер, соответствующий нормальному состоянию. Следует отметить, что в методах неконтролируемой кластеризации реальные метки классов часто неизвестны, что ограничивает возможности оценки точности кластеризации. В таких случаях можно прибегнуть к косвенным методам оценки, например, анализу свойств сформированных кластеров.

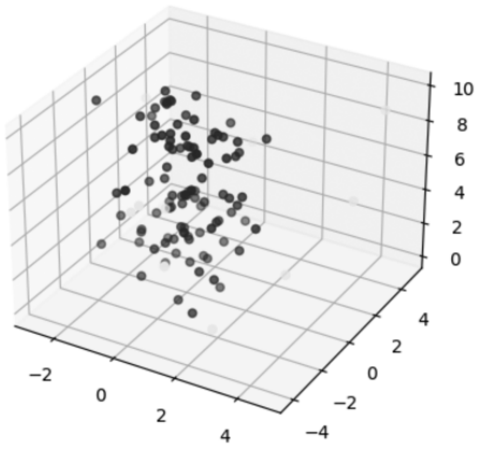


Рис. 8. Диаграмма рассеяния синтетического набора данных с тремя признаками

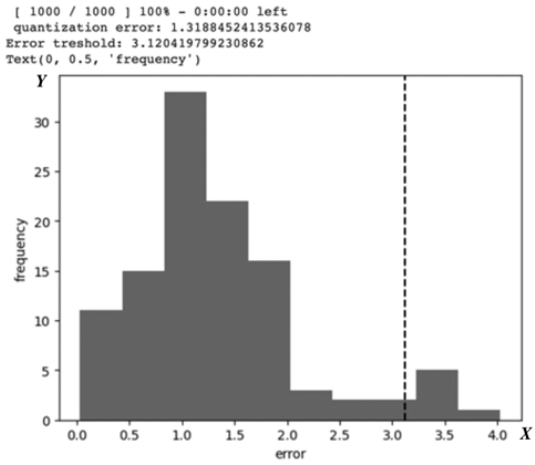


Рис. 9. Гистограмма ошибки квантования для данных с тремя признаками

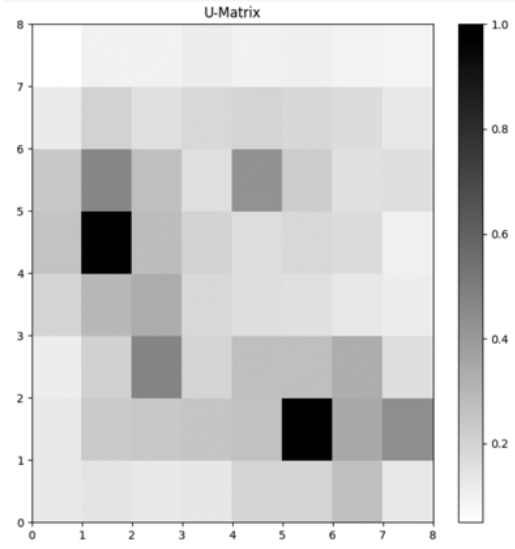


Рис. 10. Унифицированная матрица расстояний (U-matrix) для данных с тремя признаками

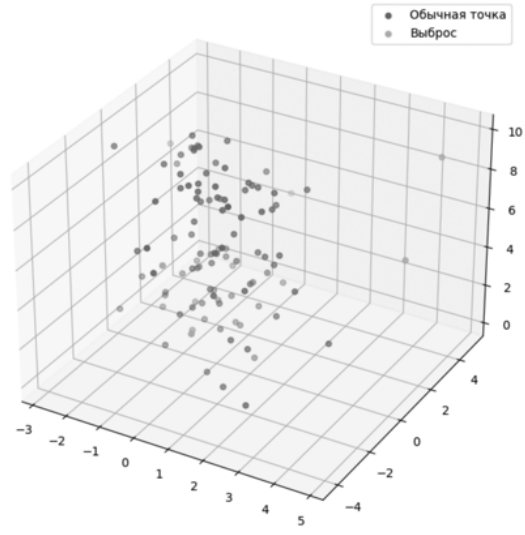


Рис. 11. Диаграмма рассеяния после обработки данных с тремя признаками картой Кохонена

Аналогичный эксперимент проводился и на более сложных данных, с увеличением признаков до трёх (рис. 8). В результате применения метода карты Кохонена на первом этапе были выделены также шесть из десяти исходных точек. При обработке синтетического трехмерного набора данных алгоритмами кластеризации на втором этапе наблюдалось, что алгоритм Isolation Forest продемонстрировал наилучшую производительность, в отличие от алгоритмов DBSCAN и OPTICS, точно выделяя представленные точки и идентифицируя их как аномальные. Но в результате одноэтапной кластеризации алгоритмом Isolation Forest точность снизилась на 7 % (рис. 13).

В условиях постоянного увеличения количества признаков (и, следовательно, размерности) обрабатываемых данных традиционные одноэтапные алгоритмы кластеризации начинают допускать ошибки, классифицируя шумовые или аномальные данные как элементы нормальных кластеров, что в свою очередь ведет к увеличению процента ложных срабатываний и снижению общей точности кластеризации (табл. 1).

В отличие от этого, двухэтапный метод кластеризации предлагает альтернативный подход, который разбивает процесс на два этапа. На первом этапе осуществляется предварительная кластеризация, которая позволяет идентифицировать основные структуры

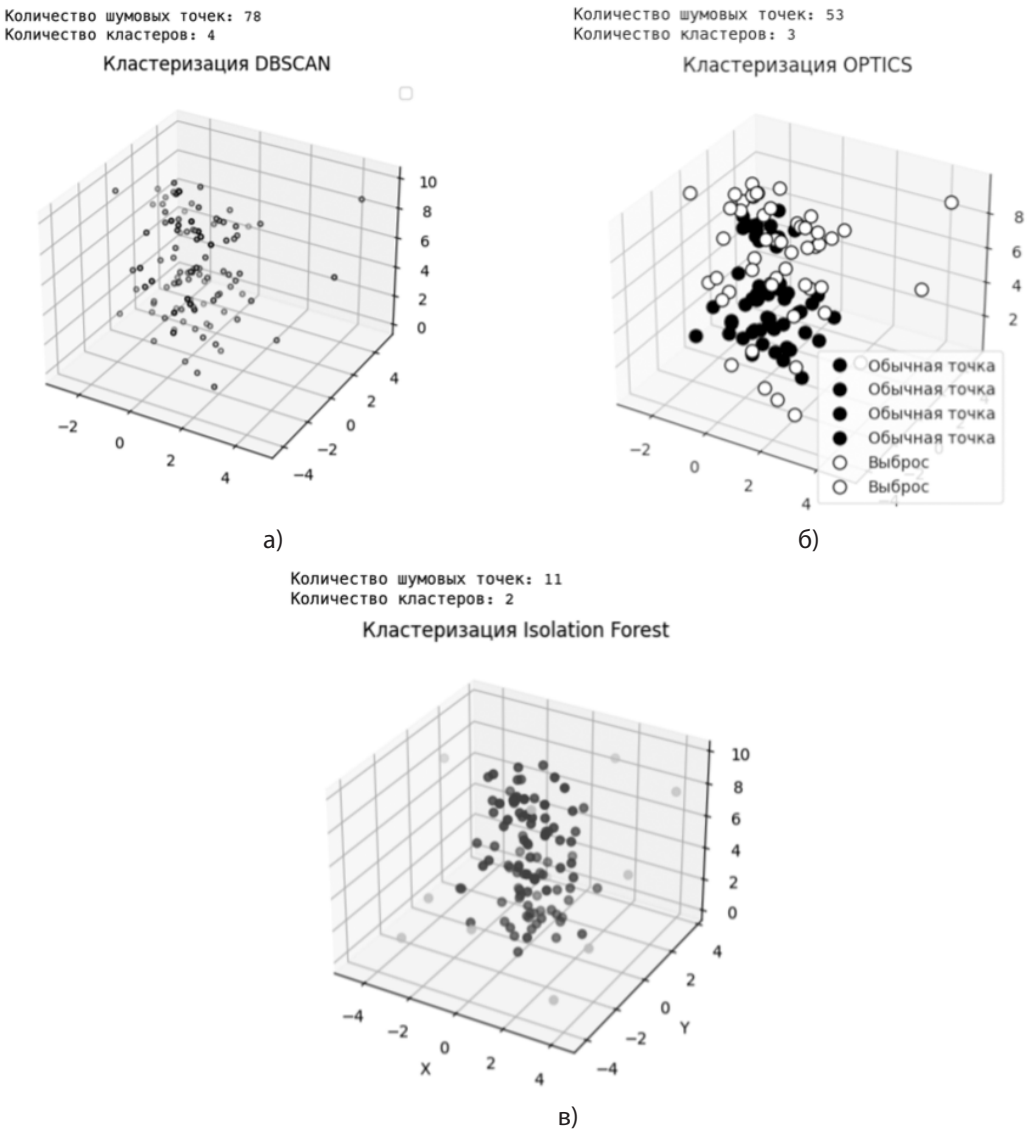


Рис. 12. Диаграммы рассеяния второго этапа кластеризации методами: DBSCAN (а); OPTICS (б) и Isolation Forest (в)

Таблица 1.

Точность кластеризации различными алгоритмами

Алгоритм	Количество признаков данных	Вид кластерной обработки	Точность
DBSCAN	2	двухэтапный	90%
OPTICS	2	двухэтапный	86%
Isolation Forest	2	двухэтапный	97%
DBSCAN	3	двухэтапный	75%
OPTICS	3	двухэтапный	72%
Isolation Forest	3	двухэтапный	97%
DBSCAN	3	одноэтапный	72%
OPTICS	3	одноэтапный	69%
Isolation Forest	3	одноэтапный	90%

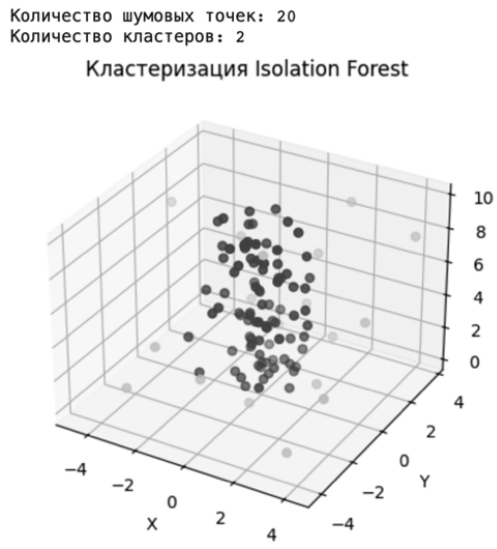


Рис. 13. Диаграмма рассеяния одноэтапной кластеризации алгоритмом Isolation Forest

в данных и определить некоторое количество шумовых элементов сразу. Этот этап также способствует сокращению объема данных, что делает возможным их последующий анализ с использованием более сложных алгоритмов кластеризации на втором этапе.

Заключение

Применение карты Кохонена на первом этапе анализа данных позволяет визуализировать и изучить их структуру, когда требуется общее представление без детального разделения на кластеры. После применения карты Кохонена для достижения более точного разделения данных на кластеры и выделения в них аномальных точек необходимо интегрировать традиционные алгоритмы кластеризации, которые проведут дополнительный анализ формируемых участков и выявят аномалии, которые могли быть упущены ранее.

Результаты, полученные на втором этапе, показали, что эффективность различных алгоритмов варьировалась в зависимости от размерности данных. Алгоритмы, такие как Isolation Forest и DBSCAN, могут требовать минимальной настройки параметров для достижения приемлемых результатов, в то время как OPTICS может требовать более глубокого понимания предварительных условий, чтобы добиться точного выделения кластеров.

Таким образом, двухэтапный метод кластеризации не только оптимизирует вычислительную нагрузку, но и улучшает конечные результаты кластеризации, обеспечивая более надежный и устойчивый анализ данных. Это особенно важно в контексте больших данных и сложных многомерных наборов, где традиционные подходы могут оказаться недостаточными.

Литература

1. Бухарев Д.А., Соколов А.Н., Рагозин А.Н. Применение иерархического кластерного анализа для кластеризации данных информационных процессов АСУ ТП, подвергающихся воздействию кибератак / Д.А. Бухарев, А.Н. Соколов, А.Н. Рагозин // Вестник УрФО. Безопасность в информационной сфере, 2023, Т. 1, №47. С. 59-68.

2. Митин Г.В., Панов А.В. Модификация алгоритма DBSCAN с использованием гибридных подходов к определению границ кластеров для обработки потоковых данных // Электронный научный журнал «ИТ-Стандарт», 2023, выпуск № 4, С. 36-57.

3. Мангутова Е.А. Обзор современных алгоритмов кластеризации данных / Мангутова Е.А., Гончаров А.С.; Томский политехнический университет, ОИТ // Молодежь и современные информационные технологии: сборник трудов XX Международной научно-практической конференции студентов, аспирантов и молодых учёных, 20-22 марта 2023 г., г. Томск. Томск: Изд-во ТПУ, 2023. С. 242-243.

4. Xia, Y., Wang D. A Survey on Density-Based Clustering Algorithms // International Journal of Computer Applications, 2015. 120(2). pp. 21-28.

5. Liu F.T., Ting K.M., Zhou Z.H. Isolation Forest // Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on. IEEE, 2008. pp. 413-422.

6. Шкодырев В.П., Обзор методов обнаружения аномалий в потоках данных / В.П. Шкодырев, К.И. Ягафаров, В.А. Баштовенко, Е.Э. Ильина; Proceedings of the Second Conference on Software Engineering and Information Management, Saint Petersburg, Russia, April 21, 2017, p. 7.

7. Xu Rui, Wunsch D.C. Survey of Clustering Algorithms. Neural Networks, // IEEE Transactions 16(3). 2005. pp. 645-678. 10.1109/TNN.2005.845141.

8. Kohonen T. Self-Organizing Maps // Berlin: Springer-Verlag. 2001. p. 502.

9. Hodge, V.J. and Austin, J. A survey of outlier detection methodologies // Artificial Intelligence Review, (2004). 22 (2). pp. 85-126.

10. Hoaglin D.C. Volume 16: How to Detect and Handle Outliers. 2013. p. 77.

11. Гундина М.А., Богдан П.С., Юхновская О.В. Особенности процесса определения количества аномальных значений при обработке измерительной информации // Вестник Белорусско-Российского университета. 2024, №2(83). URL: https://cyberleninka.ru/article/n/osobennosti-protsessa-opredeleniya-kolichestva-anomalnyh-znacheniy-pri-obrabotke-izmeritelnoy-informatsii (дата обращения: 31.10.2024).

12. Шадрин А.В. Визуализация трехмерных карт Кохонена с гексагональной решеткой // Цифровая обработка сигналов. 2015, №2. С. 23-27.

References

1. Buharev D.A., Sokolov A.N., Ragozin A.N. Primenenie ierarkhicheskogo klasternogo analiza dlya klasterizacii dannykh informacionnykh processov ASU TP, podvergayushchikhsya vozdejstviyu kiberatak / D.A. Bukharev, A.N. Sokolov, A.N. Ragozin // Vestnik URFO. Bezopasnost' v informacionnoj sfere, 2023, T. 1, № 47. S. 59-68.

2. Mitin G.V., Panov A.V. Modifikaciya algoritma DBSCAN s ispol'zovaniem gibridnykh podkhodov k opredeleniyu granic klasterov dlya obrabotki potokovykh dannykh // Ehlektronnyj nauchnyj zhurnal «IT-StandarT», 2023, vypusk № 4, S. 36-57.

3. Mangutova E.A. Obzor sovremennykh algoritmov klasterizacii dannykh / Mangutova E.A., Goncharov A.S.; Tomskij politekhnicheskij universitet, OIT // Molodezh' i sovremennye informacionnye tekhnologii: sbornik trudov XX Mezhdunarodnoj nauchno-prakticheskoj konferencii studentov, aspirantov i molodykh uchyonykh, 20-22 marta 2023 g., g. Tomsk. — Tomsk: Izd-vo TPU, 2023. — S. 242-243.9. Liu F.T., Ting K.M., Zhou Z.H. Isolation Forest // Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on. IEEE, 2008. pp. 413-422.

4. Xia, Y., Wang D. A Survey on Density-Based Clustering Algorithms // International Journal of Computer Applications, 2015. 120(2). pp. 21-28.

5. Liu F.T., Ting K.M., Zhou Z.H. Isolation Forest // Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on. IEEE, 2008. pp. 413-422.

6. Shkodyrev V.P., Obzor metodov obnaruzheniya anomalij v potokakh dannykh / V.P. Shkodyrev, K.I. Yagafarov, V.A. Bashtovenko, E.EH. Il'ina; Proceedings of the Second Conference on Software Engineering and Information Management, Saint Petersburg, Russia, April 21, 2017, p. 7.

7. Xu Rui, Wunsch D.C. Survey of Clustering Algorithms. Neural Networks, // IEEE Transactions 16(3). 2005. pp. 645-678. 10.1109/TNN.2005.845141.

8. Kohonen T. Self-Organizing Maps // Berlin: Springer-Verlag. 2001. p. 502.

9. Hodge, V.J. and Austin, J. A survey of outlier detection methodologies // Artificial Intelligence Review, (2004). 22 (2). pp. 85-126.

10. Hoaglin D.C. Volume 16: How to Detect and Handle Outliers. 2013. p. 77.

11. Gundina M.A., Bogdan P.S., Yuhnovskaya O.V. Osobennosti processa opredeleniya kolichestva anomal'nykh znachenij pri obrabotke izmeritel'noj informacii // Vestnik Belorussko-Rossijskogo universiteta. 2024. №2 (83). URL: https://cyberleninka.ru/article/n/osobennosti-protsessa-opredeleniya-kolichestva-anomalnyh-znacheniy-pri-obrabotke-izmeritelnoy-informatsii (data obrashcheniya: 31.10.2024).

12. Shadrin A.V. Vizualizaciya trekhmernykh kart Kokhonena s geksagonal'noj reshetkoj // Cifrovaya obrabotka signalov. 2015, №2. S. 23-27.

Плетенкова Анастасия Дмитриевна, аспирант кафедры защита информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: pletenkovaad@susu.ru

Соколов Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой защиты информации ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: sokolovan@susu.ru

Pletenkova Anastasia Dmitrievna, post-graduate student of the Information Security, Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: pletenkovaad@susu.ru

Sokolov Alexander Nikolayevich, Candidate of Technical Sciences, Associate Professor, Head of Department of Information Security, Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 454080, Chelyabinsk, Lenina avenue, 76. E-mail: sokolovan@susu.ru.

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 № 4(54) / 2024
Подписано в печать 24.12.2024. Дата выхода в свет 26.12.2024.
Формат 70×108 1/16. Печать цифровая. Усл.-печ. л. 5,60. Тираж 50 экз.
Заказ 406/509.
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. 4(54) / 2024
Signed to print 24.12.2024. Date of publication of the 26.12.2024.
Format 70×108 1/16. Screen printing. Conventional printed sheet 5,60. Circulation – 50 issues.
Order 406/509.
Open price.

Printed in the printing house of the Publishing Center of FGAOU VO «SUSU (NIU)».
SUSU, Publishing Center, 76, Lenina Str., Chelyabinsk, 454080