DOI: 10.51790/2712-9942-2024-5-2-10

ПРОВЕРКА ДАННЫХ НА МОДИФИКАЦИЮ В КОРПОРАТИВНОЙ СЕТИ ПРИ ПОМОЩИ КОЭФФИЦИЕНТА ДОСТОВЕРНОСТИ

В. А. Жебель[1,а], **А. И. Солдатов**[2,б]

[1] *Сургутский государственный университет, г. Сургут, Российская Федерация*

[2] *Томский государственный университет систем управления и радиоэлектроники, г. Томск, Российская Федерация*

[а] *ORCID: http://orcid.org/0009-0003-9625-5250, ✍ vladzhebel@yandex.ru*

[б] *ORCID: https://orcid.org/0000-0003-1892-1644, asoldatof@mail.ru*

Аннотация: в статье представлены исследования в рамках защиты сетевого трафика от атак на целостность, также предложен метод определения достоверности передаваемых данных по компьютерной сети предприятия. Предложенный метод основан на использовании проверочных пакетов для анализа активности сети предприятия. В работе приведено детальное описание инструментов, используемых для реализации данного метода. Эти инструменты включают в себя программное обеспечение, разработанное на языке программирования Python 3 для серверной части с применением сетевых библиотек Scapy и Socket, а также клиентскую программу для рабочей станции. В статье также рассматриваются выявленные проблемы, возникшие в процессе исследования коэффициента достоверности данных. Предложен подход для их решения при расчете данного коэффициента. Кроме того, описаны основные аспекты качественных и количественных критериев для оценки коэффициента достоверности и рассмотрена методика их расчета. В заключении приведены обобщенные результаты исследования, оценка эффективности используемого программного обеспечения, а также представлены схемы и алгоритмы функционирования предложенного метода в сетевой среде.

Ключевые слова: анализ данных, сетевой трафик, коэффициент достоверности, компьютерная сеть, качественные и количественные критерии.

Для цитирования: Жебель В. А., Солдатов А. И. Проверка данных на модификацию в корпоративной сети при помощи коэффициента достоверности. *Успехи кибернетики*. 2024;5(2):90–96. DOI: 10.51790/2712-9942-2024-5-2-10.

Поступила в редакцию: 06.04.2024. *В окончательном варианте: 09.06.2024.*

DETECTION OF DATA MODIFICATIONS IN CORPORATE NETWORKS USING A CONFIDENCE COEFFICIENT

V. A. Zhebel[1,a], **A. I. Soldatov**[2,b]

[1] *Surgut State University, Surgut, Russian Federation*

[2] *Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation*

[а] *ORCID: http://orcid.org/0009-0003-9625-5250, ✍ vladzhebel@yandex.ru*

[б] *ORCID: https://orcid.org/0000-0003-1892-1644, asoldatof@mail.ru*

Abstract: this article presents research focused on protecting network traffic from integrity attacks and introduces a method for determining the reliability of transmitted data within corporate networks. The proposed method utilizes verification packets to analyze network activity. A comprehensive description of the tools used to implement this method is provided, including the server side software developed in Python 3, leveraging the Scapy and Socket network libraries, and a workstation client. The article also addresses the challenges encountered during the study of the data reliability coefficient and suggests solutions for these issues. Additionally, the article outlines the main aspects of qualitative and quantitative criteria for assessing the reliability coefficient and details their estimation methods. The conclusion presents the overall results of the study, evaluates the performance of the software used, and describes possible applications of the proposed method in network environments.

Keywords: data analysis, network traffic, reliability coefficient, computer network, qualitative and quantitative criteria.

Cite this article: Zhebel V. A., Soldatov A. I. Detection of Data Modifications in Corporate Networks Using a Confidence Coefficient. *Russian Journal of Cybernetics*. 2024;5(2):90–96. DOI: 10.51790/2712-9942-2024-5-2-10.

Original article submitted: 06.04.2024. *Revision submitted: 09.06.2024*.

Введение

В настоящее время участились атаки на компании Российской Федерации, данные атаки направлены как на большие компании и их ресурсы, такие как vk.com, так и на государственные ресурсы, такие как www.gosuslugi.ru. Также более мелкие компании и их ресурсы периодически подвергаются атакам злоумышленников с целью наживы либо выявления конфиденциальной информации и нарушения основной деятельности.

Согласно статистике, представленной компанией Positive Technologies, в 2022 году распределение атак было следующим: госучреждения — 17%, медицинские учреждения — 9%, промышленность — 9%, наука и образование — 7%, IT-компании — 6%, сфера услуг — 5%, финансовые организации — 4%, другие отрасли — 22%, без привязки к отрасли — 21%.

Из представленной статистики видно, что 43% атак лишены четкой идентификации. Злоумышленники преследовали разнообразные цели, такие как утечка конфиденциальной информации, нарушение основной деятельности, причинение ущерба интересам государства, нанесение прямых финансовых потерь, использование ресурсов компаний или частных лиц для проведения атак, а также другие. Объекты атак включают компьютеры, серверы, сетевое оборудование, людей, веб-ресурсы, мобильные устройства и прочее. При осуществлении атак использовались различные методы, такие как вредоносное программное обеспечение, социальная инженерия, эксплуатация существующих уязвимостей, компрометация учетных данных и цепочек поставки данных, а также доверенных каналов связи. Для реализации этих методов злоумышленники прибегали к различным видам вредоносного программного обеспечения, включая шифровальщики, ПО для удаленного управления, загрузчики, шпионское ПО, майнеры, банковские трояны, ПО для удаления данных и рекламное ПО. Это программное обеспечение чаще всего распространялось через электронную почту, веб-сайты, социальные сети, мессенджеры, а также путем компрометации персональных компьютеров, серверов и сетевого оборудования [1, 2].

Для защиты передаваемых по компьютерной сети данных различные исследователи и компании предлагают разнообразные методы решения. Первым и наиболее очевидным подходом является применение шифрования [3]. Протоколы шифрования существенно усилили безопасность сетевого трафика и широко используются в течении длительного времени. Тем не менее, появление новых вычислительных устройств, способных расшифровывать трафик, а также разработка новых программ и нейронных сетей, которые могут также осуществлять расшифровку данных, приводят к необходимости постоянного совершенствования методов защиты. Следует отметить также создание квантовых компьютеров, обладающих способностью быстро расшифровывать сетевой трафик. Дальнейшее развитие систем шифрования связано с увеличением длины шифровальных ключей. Например, для протокола SSH безопасным считался ключ длиной 512 бит, затем — 1024 бит, затем — 2048 бит, и в настоящее время производители оборудования рекомендуют ключи длиной 4096 бит. Этот процесс свидетельствует о необходимости постоянного развития методов обеспечения безопасности данных. Вторым методом защиты данных являются сетевые брандмауэры [4, 5], также известные как межсетевые экраны. Эти устройства способны блокировать входящий трафик и разрешать только тот, который запрашивается пользователями. Брандмауэры могут быть аппаратными или программными, интегрированными в устройства маршрутизации. Однако они не обеспечивают защиту от подмененного трафика; злоумышленник, модифицируя передаваемый от сервера трафик, может обойти эту защиту. Кроме того, если злоумышленник находится между жертвой и сервером, он может перехватывать ответы от сервера, модифицировать трафик и направлять его к устройству жертвы [1, 6].

В связи с разнообразием целей и объектов атак мы предлагаем метод защиты, направленный на выявление атак или модификации сетевого трафика корпоративных клиентов. Метод заключатся в пересылке проверочных пакетов в количестве 10 штук в рамках сетевого протокола IPv4 с одной рабочей станции на другую, где последняя осуществляет суммирование маячков для определения коэффициента достоверности данных (КДД), величина которого характеризует, был ли модифицирован или изменен трафик [7].

В предыдущих работах [7, 8] нами были отобраны следующие протоколы и характеристики

их полей: длина кадра (LF), длина пакета (LP), длина сегмента (LS), поле пакета DSCP (DSCP), поле пакета Flag (флаги) (F), поля пакета TTL для IPv4 и Hop Limit для IPv6 (T и H), поле сегмента Urgent протокол TCP (U), поле сегмента Control bits протокол TCP (CB), поле сегмента Options протокол TCP (O), поле сегмента Sequence Number протокол TCP (SN), поле сегмента Acknowledgement Number (AN) и информация с верхнего уровня (уровень приложения) — Data. После этого возник вопрос об установлении критериев, в частности, их качественных и количественных аспектов. В ходе исследования было установлено, что характеристики полей, такие как длина кадра (LF), длина пакета (LP) и длина сегмента (LS), будут выступать в роли количественных признаков. Поля пакета TTL для IPv4 и Hop Limit для IPv6 будут сочетать в себе как качественные, так и количественные признаки. КДД определяется из выражения:

$$КДД = LF + LP + LS + DSCP + F + T + U + CB + O + SN + AN + data, \quad (1)$$

где: LF — длина кадра, LP — общая длина пакета, LS — длина сегмента, $DSCP$ — поле пакета IPv4, F — поле пакета Flag, T — поле пакета TTL для IPv4, U — поле сегмента Urgent протокол TCP, CB — поле сегмента Control bits протокол TCP, O — поле сегмента Options протокол TCP, SN — поле сегмента Sequence Number протокол TCP, AN — поле сегмента Acknowledgement Number, $data$ — данные, $КДД$ — коэффициент достоверности данных [7, 9].

Материалы и методы

По результатам предыдущих исследований [7–9] было решено ограничиться вначале исследованиями локальной вычислительной сети предприятия. В предложенном методе используется централизованное управление данными рабочих станций (ПК) с помощью сервера. На рисунке 1 изображена примерная схема корпоративной сети, включающая два сервера, персональные компьютеры (PC1-PC5), два коммутатора второго уровня и один коммутатор третьего уровня, для выхода в глобальную сеть используются два роутера по технологии виртуального сетевого адреса.

На сервере устанавливается специализированное программное обеспечение для отправки сообщений и сбора данных от рабочих станций, после чего данную информацию можно будет проанализировать. На рабочих станциях устанавливается специализированное программное обеспечение, предназначенное для мониторинга и анализа передаваемых пакетов согласно заданному алгоритму. В его основе лежит алгоритм определения коэффициента достоверности передаваемых пакетов в соответствии с формулой (1). Разработанное программное обеспечение выполнено на языке программирования Python 3 с использованием библиотек Socket и Scapy. Эти библиотеки предоставляют возможность настройки пакетов, передаваемых по сети, а также управления сеансами связи.

Принцип функционирования метода заключается в передаче пакетов на рабочую станцию, на которой закодирован один из вариантов обмена данными (рис. 2).

Данные также подлежат хранению в защищенном виде на рабочих станциях. Согласно установленной процедуре, через определенный интервал времени сервер направит на рабочие станции проверочные пакеты (рис. 3). Эти пакеты будут содержать заполненные особым образом качественные и количественные критерии.

После проверки пакетов и вычисления коэффициента достоверности рабочие станции (ПК1, ПК2, ПК3) выводят информацию о значениях достоверности данных на панели задач рабочей станции и передают ее на сервер, где она записывается в журнал (рис. 4), после чего эти данные могут быть проанализированы для принятия соответствующих решений. Важно отметить, что данные, передаваемые между сервером и рабочими станциями по сети, кодируются и шифруются для обеспечения безопасности [11].

Итак, данный метод не только позволяет выявить модификацию или изменение сетевого трафика при помощи вычисления КДД, но и обеспечивает анализ передаваемых данных, что позволяет провести анализ сетевой структуры компании. В результате использования этого метода можно выявить местоположение и наличие специализированного программного обеспечения внутри организации.

Результаты

В результате проведенных исследований были разработаны два компонента программного обеспечения на языке программирования Python 3: серверная и клиентская части. Серверная часть осуществляет отправку пакетов, в то время как клиентская принимает эти пакеты, проводит расчет ко-

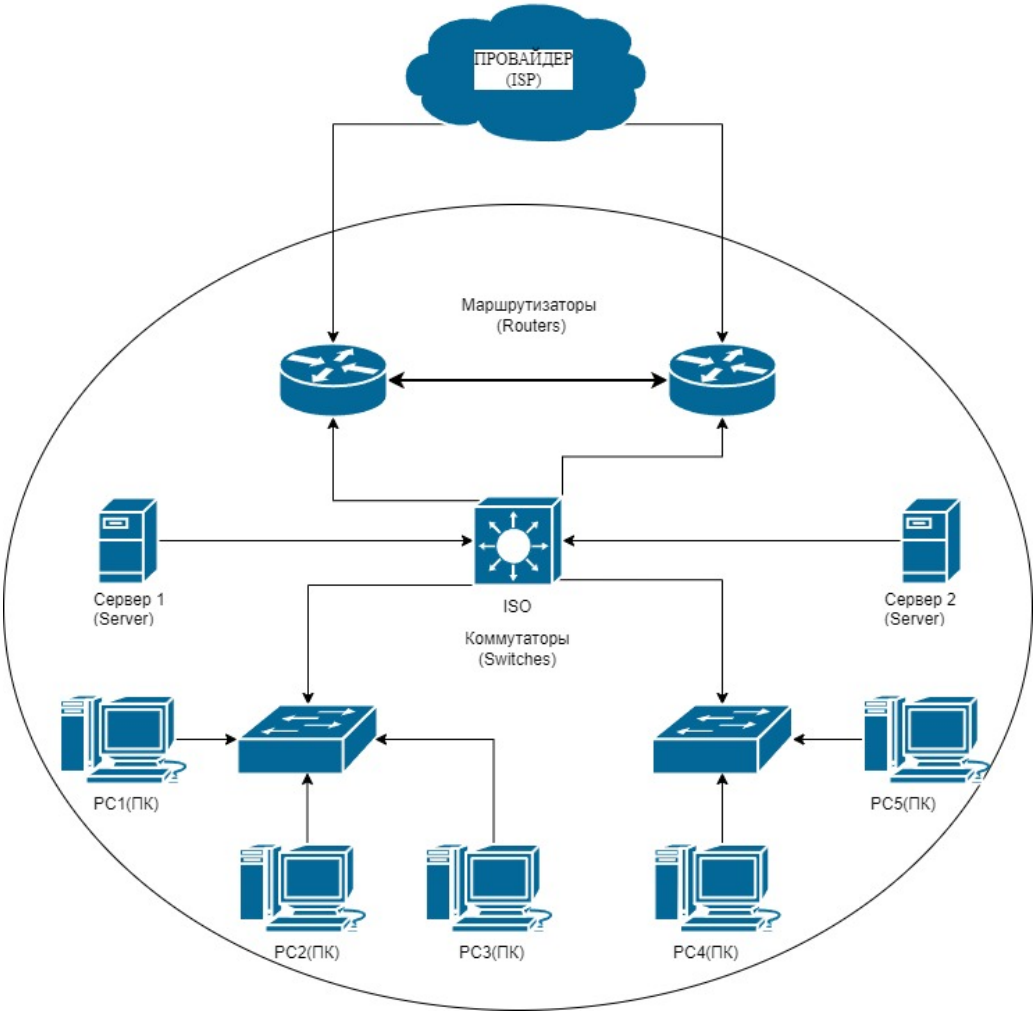

Рис. 1. *Логическая схема компьютерной сети предприятия*

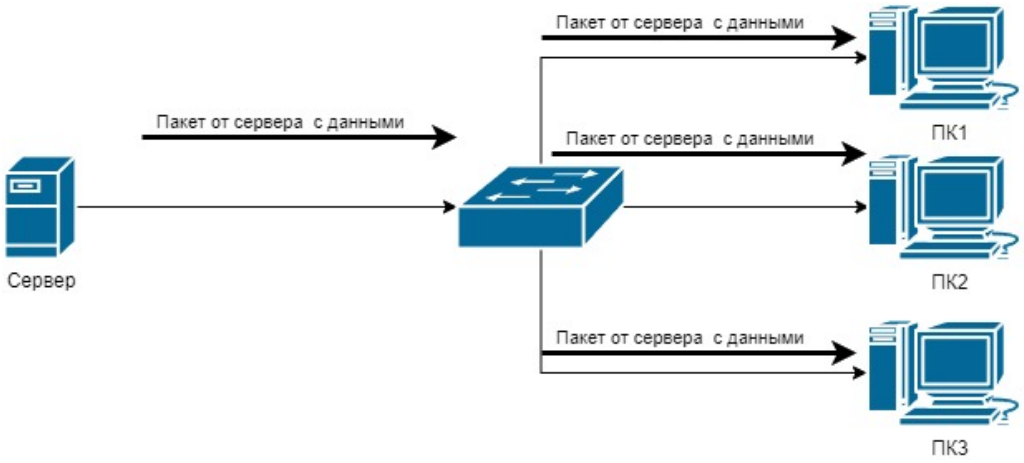

Рис. 2. *Маршруты передачи пакетов от сервера к станциям*

эффициента достоверности и передает данные обратно на сервер. При активации серверной машины проводится опрос рабочих станций в сети для установления информации о том, на каких машинах установлено клиентское программное обеспечение, и его доступности по определенному порту, например, по порту 48 888. В дальнейшем через определенные промежутки времени инициируется отправка проверочных пакетов в количестве десяти штук. В эти проверочные пакеты внедряются данные, содержащие как качественные, так и количественные признаки, закодированные в числовые значения.

Алгоритм создания проверочного пакета представлен на рисунке 5. После отправки провероч-

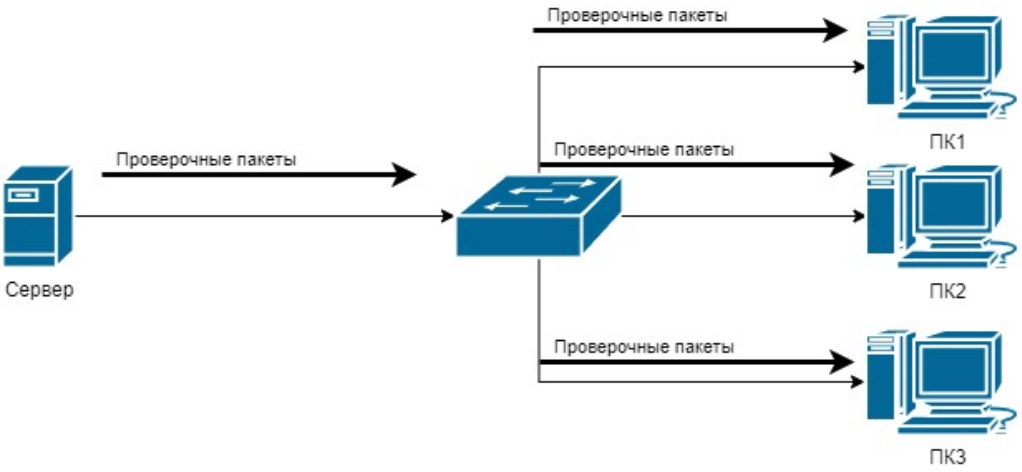

Рис. 3. *Передача проверочных пакетов от сервера к рабочим станциям*

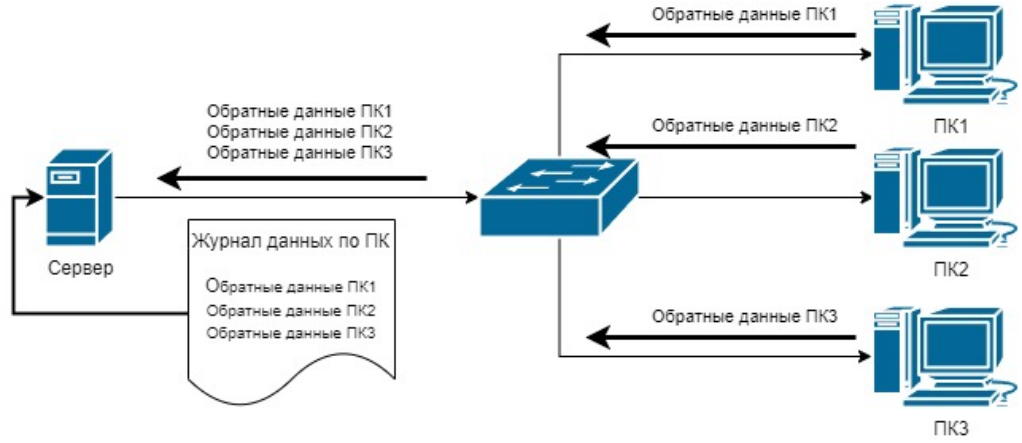

Рис. 4. *Схема обмена данными от рабочих станций к серверу*

ного пакета на рабочую станцию производится сравнение с текущим эталонным вариантом данных. Например, предположим, что у нас есть вариант данных № 7 из 1000 возможных, в котором длина сегмента составляет 70 байт, из которых 24 байта представляют заголовок, а 46 байт — сами данные. Длина пакета в этом варианте равна 90 байт, где 20 байт — это стандартный размер заголовка IPv4 пакета. Затем собирается кадр, который включает в себя 90 байт данных пакета и заголовок в 18 байт. Таким образом, общая длина кадра составляет 108 байт. Это значение больше минимального размера кадра (46 байт), но меньше максимального (1518 байт). После получения проверочных пакетов данных рабочая станция сравнивает их с ее текущим вариантом и осуществляет расчет КДД путем сравнения количественных и качественных критериев, т. е. пришедший пакет разбирается на составляющие по заголовкам протоколов IPv4, TCP и данных, после чего данные поля сравниваются с теми показаниями, что должны быть в эталонном варианте, в нашем случае № 7; если все поля соответствуют, мы ставим 1, если нет, то 0 [12, 13].

При проведении экспериментов выяснилось, что библиотека Scapy отправляет идентичные пакеты, что влияет на значения полей сегмента Sequence Number протокола TCP (SN) и Acknowledgement Number (AN). Таким образом, эти два поля заголовка сегмента можно также рассматривать как количественные критерии.

После выделения критериев качества и количества из десяти проверочных пакетов вычисляется суммарное значение каждого критерия. Пример вычисления критерия:

$$LP = \sum_{1}^{n} lp_n, \tag{2}$$

где LP — критерий длина пакета, n — количество отправленных пакетов ($n = 1..10$), lp_n — один зако-

дированный пакет данных, отправленный с сервера на рабочую станцию; если кадр дошел, величина равна 1, если нет, то 0.

Суммарный коэффициент достоверности определяется по формуле (3).

$$КДД = \frac{LP + LS + DSCP + F + T + U + CB + O + SN + AN}{100}. \quad (3)$$

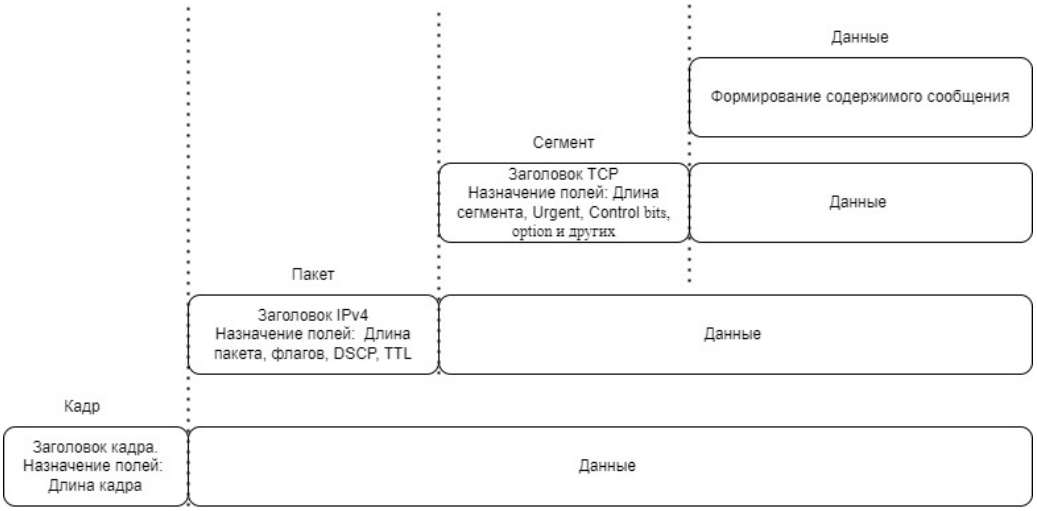

Рис. 5. *Алгоритм сборки проверочного пакета*

Минимальная величина КДД, при которой принятые пакеты считаются немодифицированными, составляет 0,9. Такая величина обусловлена возможной потерей пакетов по непредвиденному случаю. Как правило, при ответе на ICMP-запрос первый пакет не доходит, поэтому при отправке 10 пакетов получено будет 9.

В процессе исследования было принято решение убрать критерий LF — длина кадра (фрейма), так как библиотека Scapy не дает возможность напрямую считывать данный параметр (рис. 6) [14, 15].

```
>>> c=Ether(b)
>>> c
<Ether dst=00:02:15:37:a2:44 src=00:ae:f3:52:aa:d1 type=0x800 |<IP version=4L
```

Рис. 6. *Фрагмент заголовка кадра Ethernet*

Результаты расчета КДД отображаются в «tray» (область уведомлений) панели задач: зеленый цвет указывает на нормальное состояние, желтый — на возможные проблемы, а красный — на обнаруженные проблемы.

Заключение

В ходе проведенных исследований был предложен способ выявления аномалий сетевого трафика. Способ реализован в виде программы, которая состоит из нескольких ключевых компонентов. Первый компонент — серверное программное обеспечение, осуществляющее централизованное управление сетью компьютеров. Это программное обеспечение выполняет поиск рабочих станций, их синхронизацию с сервером и затем, в соответствии с определенным алгоритмом, инициирует отправку проверочных пакетов. Второй компонент устанавливается на рабочих станциях и осуществляет обмен данными по определенному порту. Он устанавливает соединение с сервером, ожидает получения проверочных пакетов, сравнивает их с назначенными вариантами от сервера, а затем проводит расчет КДД. Рабочая станция также отправляет сообщение на сервер о состоянии сети. Эти сообщения собираются в журнале, где они сортируются по рабочим станциям, а также делятся на файлы в хронологическом порядке.

Расчет коэффициента достоверности производится путем суммирования количественных и качественных критериев. Затем вычисляется основной коэффициент, который получается путем суммирования и деления на 100. Допустимое отклонение коэффициента достоверности может достигать 10% от максимального значения, при этом считается, что принятые данные не модифицированы. В настоящее время предложенный подход реализован для операционных систем семейства Windows, которые являются базой для проведения текущих и будущих исследований. Также были предприняты попытки работы с Linux-системами, но пока они находятся на начальных этапах разработки.

ЛИТЕРАТУРА

1. *Актуальные киберугрозы: итоги 2022 года*. Режим доступа: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-threatscape-2022.
2. Багдасарян Р. Х. О современных проблемах проверки достоверности данных при передаче информации и компрометации канала связи. *IX Международная научно-практическая конференция молодых ученых, посвященная 58-й годовщине полета Ю. А. Гагарина в космос: сборник научных статей.* Краснодар: Издательский Дом — Юг; 2019. С. 289–291.
3. Григорьев В. С. Проблемы распознавания зашифрованного трафика в канале связи. *Научные записки молодых исследователей*. 2017;3:43–51.
4. Сеидова И., Каратова Д. Брандмауэры: исследование методов безопасности и угроз. *Sciences of Europe.* 2022;107:137–139.
5. Коломойцев В. С. Эффективность поэтапного применения средств защиты с пересечением областей обнаружения угроз. *Программные продукты и системы*. 2018;3:557–564. DOI: 10.15827/0236-235X.031.3.557-564.
6. Буковшин В. А., Чуб П. А., Короченцев Д. А. Анализ зашифрованного сетевого трафика на основе вычисления энтропии и применения нейросетевых классификаторов. *Известия ЮФУ. Технические науки.* 2020;6:117–128. DOI: 10.18522/2311-3103-2020-6-117-128.
7. Жебель В. А., Солдатов А. И. Использование коэффициента достоверности данных для определения достоверности передаваемых данных по сети. *Успехи кибернетики.* 2023;4(2):60–67. DOI: 10.51790/2712-9942-202.
8. Жебель В. А. Достоверность данных, переданных по компьютерной сети. *International Conference on Advances in Environment Research.* 2022;3:73–78.
9. Жебель В. А. Методы и алгоритмы анализа сетевого трафика для определения коэффициента достоверности данных. *Проблемы и решения автоматизации XXI века: материалы VI Национальной научно-практической студенческой конференции*. Сургут; 2023. С. 23–27.
10. Бабенко Л. К. *Криптографическая защита информации: симметричное шифрование: учебное пособие*. М.: Юрайт; 2019. 220 с.
11. Бухарин В. В. Метод обнаружения сетевого перехвата информационного трафика информационно-телекоммуникационной сети. *Электронный журнал «Труды МАИ»*. 2012;57:1–9.
12. Багдасарян Р. Х., Багдасарян В. О. О технологии распределенной передачи данных и проблемах проверки достоверности информации по каналу связи. *Прикаспийский журнал: управление и высокие технологии.* 2021;56:48–57.
13. *Philippe BIONDI Network packet manipulation with Scapy.* Режим доступа: https://scapy.net/talks/scapy_hack.lu.pdf.
14. *Официальный сайт Scapy.* Режим доступа: https://scapy.net/.
15. Каяшев А. И., Рахман П. А., Шарипов М. И. Анализ показателей надежности локальных компьютерных сетей. *Вестник Уфимского государственного авиационного технического университета.* 2013;5:140–149.