Немчинов А.В. Методы обнаружения аномалий в информационной инфраструктуре на основе статистического анализа и корреляции трафика // Международный журнал информационных технологий и энергоэффективности.– 2026. –Т. 11 № 1(63) с. 88–94

Международный журнал информационных технологий и энергоэффективности

Сайт журнала:
http://www.openaccessscience.ru/index.php/ijcse/

УДК 004.056.57

МЕТОДЫ ОБНАРУЖЕНИЯ АНОМАЛИЙ В ИНФОРМАЦИОННОЙ ИНФРАСТРУКТУРЕ НА ОСНОВЕ СТАТИСТИЧЕСКОГО АНАЛИЗА И КОРРЕЛЯЦИИ ТРАФИКА

Немчинов А.В.
ФГБОУ ВО САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФЕССОРА М. А. БОНЧ-БРУЕВИЧА, Санкт-Петербург, Россия (193232, г. Санкт-Петербург, просп. Большевиков, 22, корп. 1), e-mail: sasha01082004@gmail.com

Статья посвящена методам обнаружения аномалий в информационной инфраструктуре на основе статистического анализа и корреляции сетевого трафика. В статье описываются факторы нормальной работы сети и ее активность, метрики важные для анализа состояния сети, какие закономерности обычно их распределения используются для выявления отклонений.

Ключевые слова: Статистический анализ, корреляция трафика, мониторинг, аномалии, обнаружение угроз, информационная инфраструктура.

METHODS FOR DETECTING ANOMALIES IN THE INFORMATION INFRASTRUCTURE BASED ON STATISTICAL ANALYSIS AND TRAFFIC CORRELATION

Nemchinov A.V.
ST. PETERSBURG STATE UNIVERSITY OF TELECOMMUNICATIONS NAMED AFTER PROFESSOR M. A. BONCH-BRUEVICH, St. Petersburg, Russia (193232, St. Petersburg, ave. Bolshevikov, 22, bldg. 1), e-mail: sasha01082004@gmail.com

The article is devoted to methods for detecting anomalies in the information infrastructure based on statistical analysis and correlation of network traffic. The article describes the factors of the normal operation of the network and its activity, metrics important for analyzing the state of the network, which patterns of their distribution are usually used to identify deviations.

Keywords: Statistical analysis, traffic correlation, monitoring, anomalies, threat detection, information infrastructure.

Введение

Современные методы обнаружения кибератак на основе шаблонов или сигнатур позволяют успешно выявлять только известные типы атак. По мере расширения сети и увеличения трафика многие неизвестные угрозы становятся невидимыми или могут обходить традиционные механизмы защиты. Поведенческие методы не всегда подходят к конкретной сети и оценивают ее характеристики, а сигнатурные методы не успевают за новыми и усовершенствованными атаками. Соответственно, необходимо внедрять подходы, которые фокусируются не на содержании трафика, а на его статистических характеристиках и отклонениях от нормального поведения.[1]

Несмотря на огромное количество информации для анализа, трафик всегда имеет свои закономерности, и они различны для каждой сети. Определение нормального поведения сети может показаться сложной задачей, но достаточно проанализировать определенный период времени и принять во внимание, какие отклонения от нормы допустимы. После этого будут известны "правила", по которым работает сеть, это динамическая информация, поэтому важно учитывать любые изменения в сети, чтобы избежать ложных срабатываний и постоянно обновлять эти "правила".

Обычный сетевой трафик характеризуется стабильными статистическими закономерностями, включая распределение объема передаваемых данных, количество активных подключений, временные интервалы между пакетами, а также использование портов и протоколов. Значительные отклонения этих параметров от типичных значений могут указывать на сбои, ошибки конфигурации или попытки вмешательства в работу сети. Например, резкое увеличение числа кратковременных подключений может указывать на сканирование портов, а изменение корреляции между объемом трафика и количеством активных узлов может указывать на аномалии маршрутизации или появление нежелательного трафика.

Для выявления таких отклонений целесообразно использовать комбинацию статистического и корреляционного анализа. [2] Статистический анализ позволяет определить границы нормального поведения сети, в то время как корреляционный анализ позволяет выявить нарушения устойчивых взаимосвязей между параметрами трафика. Совместное использование этих методов обеспечивает более полное понимание происходящих изменений и упрощает локализацию источников аномальной активности.

Практическая реализация этого подхода может быть основана на сборе статистики сетевого трафика с маршрутизаторов и других сетевых устройств, анализе ее с фиксированными интервалами и сравнении текущих значений с эталонной моделью нормального поведения. Особенно важно учитывать не только средние значения показателей, но и форму их распределения, а также взаимосвязь между различными параметрами.[3]

Для построения модели нормального поведения сети используются такие параметры, как объем данных, передаваемых за единицу времени, количество активных подключений, распределение трафика по протоколам и портам, средняя и максимальная длины пакетов, а также интервалы между пакетами. Во многих случаях эти параметры подчиняются хорошо известным статистическим законам, включая экспоненциальное распределение и распределение Пуассона, и проявляют выраженную суточную периодичность.

Для оценки отклонений используются такие методы, как вычисление среднего значения и стандартного отклонения, квантильный анализ, скользящие окна и контрольные карты. Превышение заранее определенных пороговых значений, рассчитанных на основе исторических данных, позволяет обнаружить потенциальную аномалию без анализа содержимого трафика.

Корреляционный анализ и выявление взаимосвязей

Анализа только отдельных показателей часто бывает недостаточно, так как многие атаки маскируются под обычную активность по отдельным показателям. В этом случае важную роль играет корреляционный анализ, который позволяет выявить взаимосвязи между различными параметрами трафика.

При нормальной работе сети существует устойчивая корреляция между рядом показателей. Например, увеличение объема передаваемых данных обычно сопровождается увеличением количества активных подключений, а изменение нагрузки на один сегмент сети влияет на соседние узлы. Нарушение этих связей может указывать на скрытые проблемы: туннелирование трафика, распределенные атаки или несанкционированное использование ресурсов.

Для анализа используются коэффициенты корреляции Пирсона или Спирмена, а также корреляционные матрицы, позволяющие оценить общее состояние сети. Значительное снижение или увеличение корреляции между ключевыми параметрами рассматривается как потенциальный признак аномалии и требует дополнительного анализа.

Обнаружение аномалий при конкретных типах атак

Статистический и корреляционный анализ сетевого трафика позволяет выявлять признаки ряда распространенных сетевых атак без использования сигнатурных методов. При сканировании портов, как правило, наблюдается резкое увеличение количества кратковременных подключений при относительно небольшом объеме передаваемых данных. Статистически это проявляется как сдвиг в распределении количества подключений при постоянных или незначительно изменяющихся значениях объема трафика, а также нарушение характерной корреляции между этими параметрами. Такие отклонения могут быть выявлены с помощью пороговых значений и анализа временных рядов.[4]

Атаки типа «отказ в обслуживании» (DoS и DDoS-атаки) характеризуются значительным увеличением интенсивности сетевого трафика, увеличением количества пакетов в единицу времени и уменьшением вариабельности их параметров. В таких условиях распределение объема трафика и количества подключений становится ярко выраженным асимметричным, а отдельные значения выходят за пределы доверительных интервалов, сформированных на основе модели нормального поведения сети. Статистический анализ позволяет выявлять такие выбросы на ранних стадиях атаки, до начала критического ухудшения качества услуг.

Атаки, направленные на скрытую утечку данных или туннелирование трафика, являются одними из наиболее труднодоступных для обнаружения, поскольку отдельные параметры сетевой активности могут оставаться в пределах допустимых значений. В таких случаях основным признаком аномалии является нарушение устоявшихся «правил», включая объем передаваемых данных, количество активных узлов, используемые протоколы и временные характеристики соединений. Например, увеличение исходящего трафика при постоянном количестве активных подключений или использовании нетипичного протокола может указывать на скрытую передачу данных.

Статистические методы могут выявлять аномалии, связанные с ошибками конфигурации и несанкционированным использованием сетевых ресурсов. Появление нетипичных периодических пиков активности, изменений в ежедневных профилях нагрузки или сбоев в нормальном распределении трафика по портам и службам может указывать как на технические сбои, так и на попытки замаскировать вредоносную активность.[5]

Таким образом, использование статистического и корреляционного анализа обеспечивает идентификацию широкого спектра сетевых аномалий, включая как явные, так и скрытые формы воздействий. Этот подход дополняет традиционные меры безопасности и

повышает эффективность мониторинга информационной инфраструктуры за счет выявления отклонений от нормального поведения сети.

Выявление атак типа отказа в обслуживании по интенсивности событий

На Рисунке 1 показано изменение количества событий с течением времени во время атаки типа «отказ в обслуживании». В отличие от кратковременных аномалий, характерных для сканирования портов, атаки DoS и DDoS-атак сопровождаются устойчивым и значительным увеличением интенсивности событий в течение длительного периода времени. В то же время значения временных рядов выходят за пределы доверительных интервалов, основанных на модели нормального поведения сети.

Такие аномалии могут быть обнаружены с помощью методов скользящего среднего, дисперсионного анализа и контрольных карт, что позволяет обнаруживать атаки на ранней стадии и принимать своевременные меры для защиты инфраструктуры.

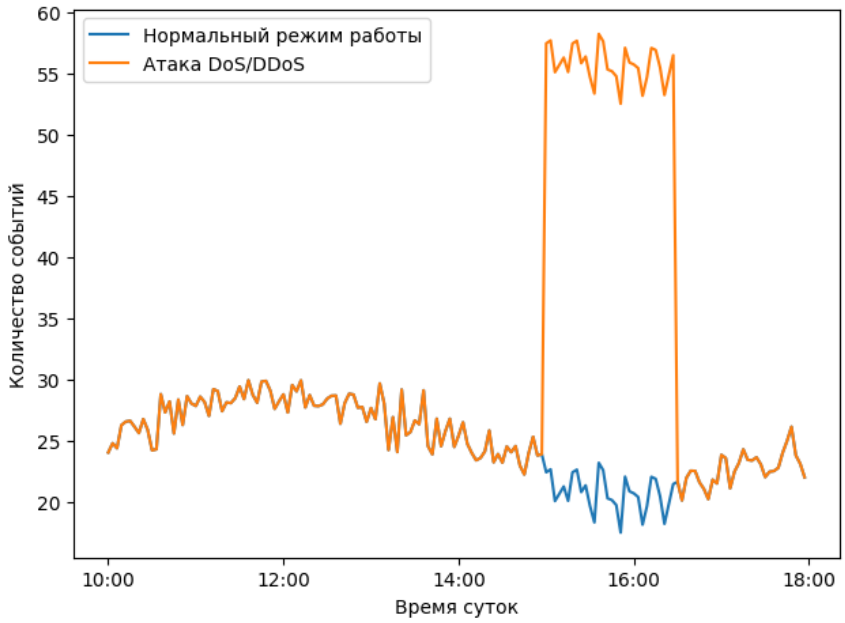

Рисунок 1 - Временной ряд количества событий

Статистический анализ распределения количества событий

Дополнительную информацию о характере сетевой активности можно получить, проанализировав распределение количества событий в единицу времени. На Рисунке 2 показано сравнение распределений количества событий в обычном режиме и в период аномальной активности. Нормальная работа сети характеризуется компактным распределением значений с ограниченной вариабельностью, тогда как во время атаки происходит сдвиг в распределении и появление значений, значительно превышающих типичные уровни.

Изменение формы распределения и увеличение дисперсии позволяют выявлять аномалии даже в тех случаях, когда средние значения показателей незначительно отличаются от нормальных. Такой подход повышает устойчивость системы обнаружения к маскировке атак под законную сетевую активность.

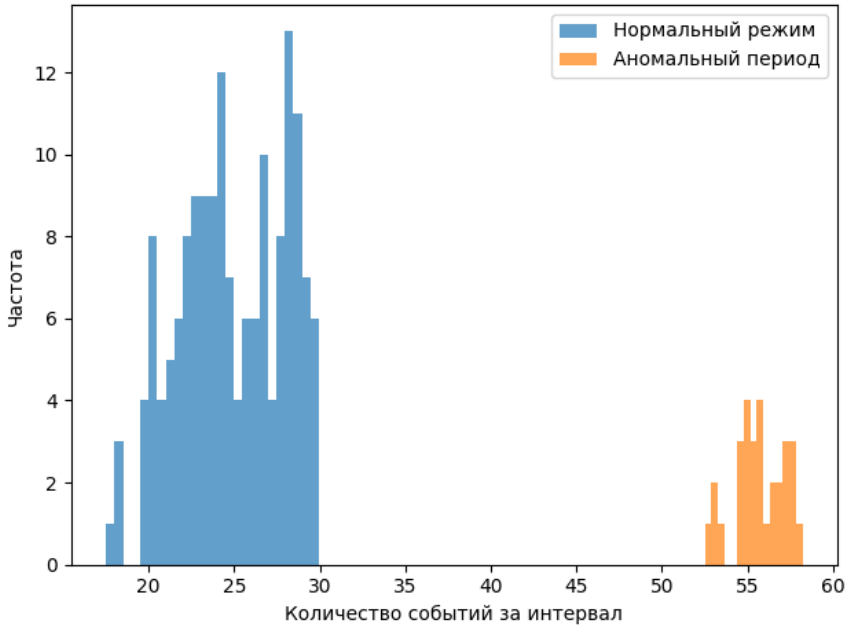

Рисунок 2 - Распределение количества событий

Применение методов в системах мониторинга

Интеграция статистических и корреляционных методов в системы мониторинга позволяет создать многоуровневый механизм обнаружения аномалий. На первом уровне регистрируются отклонения отдельных показателей от нормы, а на втором - анализируются изменения взаимосвязей между ними. Такой подход сокращает количество ложных срабатываний и повышает точность обнаружения реальных угроз.

Важным аспектом является адаптивность модели. Информационная инфраструктура со временем меняется: добавляются новые сервисы, меняется профиль нагрузки и модернизируется оборудование. Поэтому модели нормального поведения должны регулярно обновляться с учетом новых данных, а пороговые значения должны пересматриваться.

Заключение

В ходе работы рассматриваются методы обнаружения аномалий в информационной инфраструктуре, основанные на статистическом анализе и корреляции сетевого трафика и событий. Показано, что нормальное функционирование сети характеризуется стабильными статистическими закономерностями и устойчивыми корреляциями между ключевыми параметрами, такими как количество сетевых событий, интенсивность подключений и временные характеристики активности.

Анализ показывает, что отклонения от модели нормального поведения, выраженные в виде кратковременных всплесков, устойчивого превышения пороговых значений или нарушения корреляций, могут служить надежными индикаторами различных типов сетевых атак. В частности, сканирование портов проявляется в виде резкого кратковременного увеличения количества событий, атаки типа «отказ в обслуживании» характеризуются длительным увеличением интенсивности событий, а скрытая утечка данных и туннелирование трафика обнаруживаются путем нарушения стабильных взаимосвязей между параметрами сетевой активности.

Использование анализа временных рядов и распределения количества событий позволяет выявлять аномалии без анализа содержимого пакетов и использования сигнатурных методов, что особенно важно для высоконагруженных и распределенных инфраструктур. Рассмотренные подходы помогают обнаружить как явные, так и скрытые сетевые воздействия, а также помогают снизить зависимость систем мониторинга от заранее известных схем атак.

Таким образом, статистический и корреляционный анализ трафика и событий является хорошим дополнением к традиционным методам обнаружения угроз. Внедрение этих методов в системы мониторинга повышает надежность и стабильность систем, обеспечивая своевременное обнаружение аномалий и более быстрое реагирование на инциденты.

Список литературы

1. Исследование способов повышения безопасности корпоративных сетей / Н. Ф. Махмутова, Э. В. Бирих, Д. В. Сахаров [и др.] // Вестник Дагестанского государственного технического университета. Технические науки. – 2024. – Т. 51, № 3. – С. 110-116. – DOI 10.21822/2073-6185-2024-51-3-110-116. – EDN HDGBOY.
2. Развитие стандартов и руководств в сфере облачных технологий / Э. В. Бирих, Л. А. Виткова, М. В. Левин, М. В. Чмутов // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2017) : Сборник научных статей VI Международной научно-технической и научно-методической конференции. В 4-х томах, Санкт-Петербург, 01–02 марта 2017 года / Под редакцией С.В. Бачевского. Том 2. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2017. – С. 92-95. – EDN YRPZWJ.
3. Выбор инструментов динамического анализа безопасности web-приложений для задач цифровой экономики / Э. В. Бирих, А. С. Груздев, А. О. Камалова, Д. В. Сахаров // Защита информации. Инсайд. – 2024. – № 1(115). – С. 42-46. – EDN RLNHWK.
4. Разработка программного модуля для автоматизации определения уровня защищенности в ИСПДН / Э. В. Бирих, М. Д. Булова, А. А. Казанцев, А. А. Миняев // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2024) : Материалы XIII Международной научно-технической и научно-методической конференции, Санкт-Петербург, 27–28 февраля 2024 года. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2024. – С. 122-127. – EDN FBPSIL.
5. Бирих, Э. В. Современные проблемы обеспечения внутренней безопасности распределенной сети органов государственной власти / Э. В. Бирих, А. С. Гаврилов, Е. Н. Сацук // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018) : VII Международная научно-техническая и научно-методическая конференция. Сборник научных статей. В 4-х томах, Санкт-Петербург, 28 февраля – 01 2018 года / Под редакцией С.В. Бачевского. Том 1. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2018. – С. 104-107. – EDN XSUFFR.

References

1. 1. Research of methods for increasing the security of corporate networks / N. F. Makhmutova, E. V. Birikh, D. V. Sakharov [et al.] // Bulletin of the Dagestan State Technical University.

Technical sciences. - 2024. - Vol. 51, No. 3. - pp. 110-116. - DOI 10.21822/2073-6185-2024-51-3-110-116. - EDN HDGBOY.

2. Development of standards and guidelines in the field of cloud technologies / E. V. Birikh, L. A. Vitkova, M. V. Levin, M. V. Chmutov // Actual problems of infotelecommunications in science and education (APINO 2017): Collection of scientific articles of the VI International scientific-technical and scientific-methodical conference. In 4 volumes, St. Petersburg, March 1–2, 2017 / Edited by S.V. Bachevsky. Volume 2. – St. Petersburg: St. Petersburg State University of Telecommunications named after prof. M.A. Bonch-Bruevich, 2017. – pp. 92–95. – EDN YRPZWJ.
3. Selecting Tools for Dynamic Security Analysis of Web Applications for Digital Economy Tasks / E.V. Birikh, A.S. Gruzdev, A.O. Kamalova, D.V. Sakharov // Information Security. Inside. – 2024. – No. 1(115). – pp. 42–46. – EDN RLNHWK.
4. Development of a software module for automating the determination of the security level in the information system for personal data protection / E. V. Birikh, M. D. Bulova, A. A. Kazantsev, A. A. Minyaev // Actual problems of infotelecommunications in science and education (APINO 2024): Proceedings of the XIII International scientific-technical and scientific-methodical conference, St. Petersburg, February 27-28, 2024. - St. Petersburg: St. Petersburg State University of Telecommunications named after prof. M.A. Bonch-Bruevich, 2024. - pp. 122-127. - EDN FBPSIL.
5. Birikh, E. V. Modern problems of ensuring internal security of a distributed network of government bodies / E. V. Birikh, A. S. Gavrilov, E. N. Satsuk // Actual problems of infotelecommunications in science and education (APINO 2018): VII International scientific-technical and scientific-methodical conference. Collection of scientific articles. In 4 volumes, St. Petersburg, February 28 – January 01, 2018 / Edited by S. V. Bachevsky. Volume 1. - St. Petersburg: St. Petersburg State University of Telecommunications named after prof. M. A. Bonch-Bruevich, 2018. - pp. 104-107. - EDN XSUFFR.