Инженерный вестник Дона, №7 (2025)
ivdon.ru/ru/magazine/archive/n7y2025/10202

Анализ подходов к обнаружению атак нулевого дня в сетях интернета вещей

С.Ю. Рыбаков, Ф.А. Ташлыков

Московский технический университет связи и информатики, Москва

Аннотация: Злоумышленники часто используют необнаруженные уязвимости системы для проведения атак нулевого дня. Существующие традиционные системы обнаружения, основанные на методах глубокого обучения и машинного обучения, не способны эффективно справляться с новыми атаками нулевого дня. Такие атаки часто остаются неправильно классифицированными, поскольку они представляют собой новые и ранее неизвестные угрозы. Расширение сетей Интернета вещей (Internet of Things - IoT) лишь способствует росту числа таких атак. В работе анализируются подходы способные обнаруживать атаки нулевого дня в сетях IoT, основанные на неконтролируемом подходе без предварительного знания об атаках или необходимости обучения систем обнаружения вторжений (Intrusion Detection System - IDS) на заранее размеченных данных.
Ключевые слова: интернет вещей, атака нулевого дня, автокодировщик, машинное обучение, нейросеть, сетевой трафик.

Введение

Системы Интернета вещей (англ. Internet of Things – IoT) оказывают значительное влияние и трансформируют повседневную деятельность, объединяя физические объекты и деятельность человека через Интернет. Однако, как и любые другие компьютерные сети, системы IoT часто становятся мишенью злоумышленников из-за слабых мер безопасности [1]. Киберпреступники активно используют IoT-устройства для создания бот-сетей, которые впоследствии применяются для проведения масштабных кибератак, включая распределенные атаки типа «отказ в обслуживании» (англ. Distributed Denial of Service – DDoS). Количество IoT-устройств непрерывно растет и по прогнозам, к 2030 году эта цифра вырастет до 29,42 миллиарда [2]. Такая динамика увеличивает риск кибератак, особенно атак нулевого дня, которые представляют собой совершенно новые угрозы, ранее не встречавшиеся в системе. Согласно исследованиям, атаки нулевого дня ответственны за 80 % всех нарушений безопасности и наносят финансовый ущерб в среднем около 1,2 миллиона долларов на одну атаку. Эти

атаки используют ранее необнаруженные или совершенно новые уязвимости системы [3].

Широкое применение устройств IoT в различных отраслях промышленности значительно увеличило поверхность атаки и повысило риск DDoS-атак. Механизм таких атак обычно включает взлом ряда подключенных к Интернету устройств, таких как ноутбуки, IP-камеры, мобильные телефоны и другие устройства. Используя их уязвимости, злоумышленники объединяют взломанные устройства в бот-сеть, которая затем направляет огромный поток запросов на целевую систему, перегружая её. В результате легитимные пользователи теряют доступ к услугам, так как система становится недоступной. При этом жертвами становятся не только целевые системы, но и все устройства, участвующие в атаке. DDoS-атаки продолжают становиться более масштабными и сложными.

В последнее время получили развитие системы обнаружения вторжений (англ. Intrusion Detection System – IDS), основанные на машинном обучении. Эти системы обучаются на исторических данных, известных как обучающие выборки, и могут предсказывать новые сценарии на их основе. Однако они также оказываются неэффективными при обнаружении атак "нулевого дня", которые остаются незамеченными в течение длительного времени, нанося серьезный ущерб организациям. Обнаружение аномалий является одним из наиболее подходящих и эффективных методов для выявления отклонений от нормального поведения сетевых систем. Этот подход основывается на анализе сетевых данных [4-6]. Если для обучения IDS доступны размеченные данные, возможно применение контролируемых или полуконтролируемых методов. Такие подходы позволяют обучить систему, используя заранее известные аномальные и нормальные данные.

Однако в случае атак "нулевого дня" обучающие данные невозможно получить до момента выявления таких атак, что делает методы

неконтролируемого обучения единственным вариантом обучения IDS. Подобные системы способны обходить ограничения традиционных методов, поскольку кибератаки постоянно эволюционируют.

Таким образом, раннее выявление кибератак, в частности атак "нулевого дня", становится критически важной задачей для защиты информационных систем.

Алгоритмы кластеризации являются ключевыми инструментами для выявления атак "нулевого дня", что особенно важно в условиях постоянно меняющегося ландшафта киберугроз поскольку способны адаптироваться к новым угрозам, где традиционные методы обнаружения оказываются неэффективными.

В то же время отдельные алгоритмы кластеризации часто демонстрируют недостаточную эффективность при обнаружении атак "нулевого дня". Для преодоления этих ограничений применяются ансамблевые модели, которые объединяют несколько алгоритмов, что позволяет нивелировать ошибки и ограничения отдельных моделей, обеспечивая более точное и надежное обнаружение атак "нулевого дня". Таким образом, ансамблевый подход играет ключевую роль в повышении эффективности систем обнаружения вторжений.

Быстрое распознавание и анализ новых моделей атак позволяет организациям минимизировать негативные последствия, включая предотвращение потенциально катастрофических последствий и обеспечения непрерывного функционирования жизненно важных сервисов.

Методология обнаружения атак нулевого дня

Методы кластеризации наиболее подходят для обнаружения кибератак "нулевого дня" в сетевых данных, особенно в условиях отсутствия меток, необходимых для обучающих методов.

Для повышения общей производительности методов кластеризации рекомендуется объединять результаты работы нескольких алгоритмов, формируя ансамбль моделей. Такой подход позволяет компенсировать недостатки отдельных методов и повысить точность и надежность обнаружения аномалий.

Ансамблевые модели обладают значительными преимуществами в задачах неконтролируемой классификации, поскольку объединяют результаты работы множества алгоритмов кластеризации. Объединяя сильные стороны каждого алгоритма и компенсируя их недостатки, ансамблевые модели минимизируют ошибки, характерные для отдельных методов, а также противодействуют возможным эффектам переобучения. Кроме того, они обеспечивают более стабильные и последовательные результаты, особенно при работе со сложными и высокоразмерными наборами данных. Благодаря способности извлекать преимущества из разных подходов, ансамблевые модели не только повышают точность, но и улучшают понимание структуры и закономерностей в анализируемых данных.

Сокращение числа признаков

Сокращение числа признаков не только снижает вычислительную нагрузку на методы кластеризации, но и значительно повышает их эффективность, особенно при работе с большими наборами данных. Процесс сокращения признаков позволяет выделить наиболее информативные компоненты данных. Такие методы, как случайная гауссовская проекция (англ. Gaussian Random Projection), удаляют зашумленные или избыточные признаки, что улучшает качество анализа. Таким образом, снижение размерности данных является первым шагом в разработке модели для обнаружения атак, что способствует улучшению производительности алгоритмов классификации.

Благодаря множеству преимуществ по сравнению с анализом главных компонент и автокодировщиками, случайная гауссовская проекция занимает особое место в задачах неконтролируемой классификации для уменьшения числа признаков.

В отличие от метода главных компонентов и автокодировщиков, случайная гауссовская проекция обладает высокой масштабируемостью благодаря линейному росту вычислительной сложности. Это делает его особенно полезным при работе с большими наборами данных, где другие методы сталкиваются с вычислительными ограничениями. Кроме того, случайная гауссовская проекция сохраняет парные расстояния между точками данных, что критично для задач неконтролируемой классификации, так как это позволяет точно выделять кластеры, независимо от наличия размеченных данных.

В отличие от автокодировщика, случайная гауссовская проекция не требует этапа обучения, поскольку работает исключительно с признаками входных данных. Это делает его идеальным для ситуаций, когда обучающие данные отсутствуют, как в случае с DDoS-атаками нулевого дня. Отсутствие этапа реконструкции позволяет случайной гауссовской проекции избегать проблем, связанных с потерей информации или искажением, часто встречающихся в алгоритмах автокодировщика. Случайная гауссовская последовательность является более предпочтительным вариантом для задач сокращения признаков в неконтролируемой классификации благодаря простоте реализации, экономичности вычислений и способности справляться с шумами. Это особенно актуально для задач обнаружения атак "нулевого дня" в сетях с большими наборами данных, где приоритет отдается высокой скорости вычислений без ущерба для точности кластеризации.

Основой случайной гауссовской проекции является лемма Джонсона-Линденштрауса, которая позволяет преобразовывать высокоразмерные

данные в пространство с пониженной размерностью, сохраняя расстояния между двумя точками данных.

Лемма **Джонсона-Линденштрауса** утверждает, что множество из n точек многомерного пространства можно отобразить в пространство гораздо меньшей размерности n таким образом, что расстояния между точками останутся почти без изменений. При этом такое отображение можно найти среди ортогональных проекций. Лемма **Джонсона-Линденштрауса** позволяет сжимать данные, представленные точками многомерного пространства, и, что более важно, сократить размерность данных без существенной потери информации.

Это гарантирует, что структура кластеров в данных сохраняется даже в пространстве с уменьшенным числом признаков и сниженной вычислительной сложностью [7]. Согласно лемме **Джонсона-Линденштрауса**, расстояние между двумя точками данных сохраняется с точностью до произвольно малого коэффициента *eps*(ϵ). Таким образом, выборки $X = [x_1,x_2,.....,x_n]$ из пространства $R^o \to R^d$, без потери ключевых взаимосвязей в данных:

$$(1-\tau)\left\|x_i - x_j\right\|_2^2 \le \left\|(x_i)-(x_j)\right\|_2^2 \le (1+\tau)\left\|x_i - x_j\right\|_2^2 \qquad (1)$$

где $f(x_i)$ и $f(x_j)$ - преобразованные пространства признаков, а x_i и x_j - исходные признаки наборов данных. Уменьшенное пространство признаков должно удовлетворять условию $d \ge \frac{\log(n)}{\epsilon^2}$. Случайная гауссовская проекция сводит признаки к матрице $R^d = \frac{1}{\sqrt{d}} R^T(X)$ из R^d наборов данных, R^T - случайная матрица, полученная случайной гауссовской проекцией.

Классификатор на основе ансамбля

Методология ансамблевой модели, представлена на рисунке 1. На первом этапе выполняется предварительная обработка наборов данных. Эти данные проходят этап выбора признаков, который направлен на уменьшение размерности набора данных, что позволяет сократить вычислительные затраты и повысить эффективность работы модели, сохраняя при этом значимую информацию для кластеризации и обнаружения аномалий.

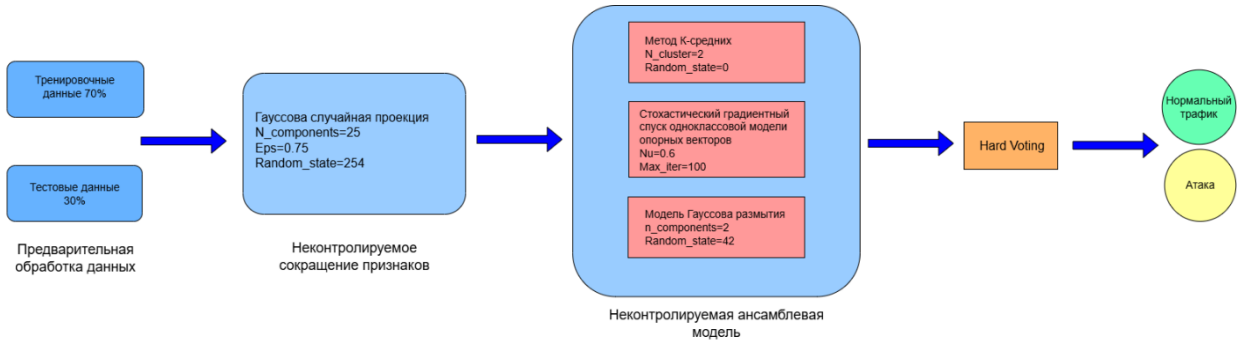

Рис. 1. – Методология, используемая в предлагаемой модели

Для классификации данных как нормальных или атакующих предлагается использовалась ансамблевую модель, объединяющую три различных классификатора, основанных на неконтролируемом обучении. В состав модели входили метод k-средних (англ. k-means), одно-классовая машина опорных векторов (англ. One-Class Support Vector Machines - OCSVM) и модель гауссовой смеси (англ. Gaussian Mixture Model). Их совместное использование позволяет повысить точность работы общей модели, несмотря на отсутствие данных с метками [8]. Важно отметить, что производительность всех используемых алгоритмов кластеризации напрямую зависит от выбранных значений гиперпараметров. Некорректная настройка гиперпараметров может значительно снизить эффективность алгоритмов, что подчеркивает важность оптимального выбора этих значений для достижения наилучших результатов.

Метод k-средних: Метод k-средних [9] — один из наиболее популярных алгоритмов кластеризации, основанных на неконтролируемом обучении. Применение метода k-средних демонстрирует высокую скорость, масштабируемость и адаптивность, что делает его особенно эффективным для обнаружения атак "нулевого дня". Алгоритм позволяет быстро выявлять аномалии в сетевом поведении, что помогает идентифицировать угрозы на ранних стадиях. Его способность обнаруживать новые модели атак в режиме реального времени делает его отличным инструментом для борьбы с DDoS-атаками "нулевого дня".

Для набора данных с d-признаками и заданным количеством кластеров c, метод k-средних находит оптимальные центры кластеров, минимизируя сумму квадратов расстояний от всех точек данных до ближайших к ним центров. Другими словами, метод k-средних использует базовую структуру данных для поиска кластеров в пределах точек данных [10]. Процесс работы алгоритма заключается в присвоении каждой точки данных ближайшего центра кластера, который служит его представителем. Таким образом, метод k-средних стремится разделить данные на k групп (кластеров) [11]. Например, алгоритм разделяет n наблюдений на k кластеров $C_1,C_2,...,C_k$, где каждое наблюдение из набора $(X_1,X_2,...,X_n)$ представляет собой вектор размерности d. Основная цель алгоритма — минимизировать разброс внутри кластеров, что достигается уменьшением суммы квадратов внутри кластера. С математической точки зрения задача заключается в определении:

$$\underset{S}{argmin} \sum_{i=1}^{k} \sum_{x \in S_i} \left\| x - u_i \right\|^2 = \underset{S}{argmin} \sum_{i=1}^{k} |S_i| Var(S_i) \quad (2)$$

где u_i - среднее значение точек данных в S_i, а для нахождения расстояния оно рассматривается как центр кластера и оптимизируется на каждой итерации.

В кластере минимизация парного квадратичного отклонения оценивается, как:

$$\underset{S}{argmin} \sum_{i=1}^{k} \left(\frac{1}{2|S_i|} \right) \sum \text{x},\ y_i \left\| x - y \right\|^2 \tag{3}$$

где $\sum x,\ y_i \left\| x - y \right\|^2 = \sum x \in \mathrm{S_i} \left(x - u_i \right)^T \left(u_i - y \right)$

SGD-OCSVM

В работе [12] для обнаружения выбросов в данных был использован метод одно-классовой машины опорных векторов, который хорошо справляется с задачей выявления аномалий, что делает его особенно подходящим для обнаружения новых атак "нулевого дня", которые выбиваются из привычных паттернов сетевого поведения. Неконтролируемый характер алгоритма и использование стохастического градиентного спуска (англ. Stochastic Gradient Descent - SGD) делают его особенно эффективным для обнаружения новых форм кибератак в режиме реального времени. Для достижения наилучших результатов алгоритм требует тщательной настройки гиперпараметров в зависимости от характеристик используемых наборов данных.

Одно-классовая машина опорных векторов обнаруживает аномалии, анализируя границу принятия решения. Стохастический градиентный спуск - это техника оптимизации при которой случайным образом используются только несколько образцов в одной партии, а не все образцы данных.

Для этого данные проецируются в пространство более высокой размерности с использованием функции ядра. Этот подход позволяет выявлять выбросы, которые отклоняются от нормального поведения, путем построения гиперплоскости, максимизирующей маржу между данными и началом координат в новом пространстве признаков [13]. Основным ограничением одно-классовой машины опорных векторов является её низкая эффективность при недостаточном количестве данных. Алгоритм может

плохо улавливать характеристики нормальных данных, что снижает его точность в обнаружении выбросов. Метод использует функции ядра для отображения исходных данных в пространство признаков, где поиск гиперплоскости осуществляется с помощью следующей оптимизационной задачи:

$$\left(\min_{\omega,\rho,\varepsilon}\right)\frac{1}{2}\|\omega\|^2+\frac{1}{vm}\sum_{i=1}^{m}\xi_i-\rho \quad (4)$$

при условии $\left(w,\phi\left(X_i\right)\right)\geq\rho-\xi_i,\ \xi_i\geq 0,\ i=1,\ldots\ldots m$, где m - общее пространство исходных признаков, v - верхний предел доли аутлайнеров и нижний предел доли векторов поддержки, X_i - i-я обучающая выборка, w - вектор гиперплоскости, $\phi()$ - карта признаков, ξ - слабые переменные, ρ - порог или смещение.

Окончательная функция принятия решения становится:

$f\left(x\right)=\sum_{alpha_i>0}a_i k\left(x_i,x_j\right)-\rho$, на основании которой любая точка данных классифицируется как промах, как атакующие данные, если $f(x)<0$ SGD-OCSVM использует стохастический градиентный спуск при применении алгоритма одно-классовой машины опорных векторов к точкам данных.

Модель гауссовской смеси: Модель гауссовской смеси — это вероятностная модель, основанная на оценке плотности, применяемая в методах неконтролируемой кластеризации и основанная на оценке плотности распределения вероятности. Она схожа с методом k-средних в аспекте анализа данных, однако вместо подхода, основанного на расстоянии, использует вероятностный подход.

Преимуществом модели гауссовской смеси является его способность моделировать сложные распределения данных, включая мультимодальные паттерны, которые часто характерные для поведения атак. Этот алгоритм

гибко отражает различные уровни сложности распределений, что делает его особенно полезным для анализа данных, содержащих сложные и неоднородные структуры. Одной из ключевых особенностей модели гауссовской смеси является мягкая кластеризация, позволяющая точкам данных принадлежать сразу нескольким кластерам с разной вероятностью. Такой подход помогает выявлять незначительные вариации в данных, которые могут служить индикаторами атак "нулевого дня". Однако модель гауссовской смеси может испытывать трудности в условиях высокоразмерных данных или в ситуациях, когда кластеры плохо разделены. Первым шагом в алгоритме модели гауссовской смеси является определение оптимального количества кластеров в наборе данных. Для этого используется байесовский информационный критерий (Bayesian information criterion - BIC) [14]. Определение числа кластеров может быть вычислительно затратным, однако в задаче обнаружения атак количество кластеров обычно фиксируется равным двум: один кластер соответствует данным с атакой, другой — данным без атаки.

Модель гауссовской смеси можно рассматривать как гибридную модель, состоящую из нескольких одномерных гауссовских распределений. Вероятностное распределение *P(x)* в модели гауссовской смеси оценивается как взвешенная сумма этих гауссовских моделей и выражается следующим образом:

$$P(x) = \sum_{k=1}^{K} \Pi_k N\left(x \middle| \mu, \sigma_k\right) \quad (6)$$

где для k-й подмодели π_k - вес вероятности с условием $\sum_{k=1}^{K} \Pi_k = 1; N\left(x \middle| \mu, \sigma_k\right)$ - соответствующее гауссовское распределение, μ_k - среднее значение и σ_k - ковариация. Наибольшая точность достигается при $n_{компонент}$ 2 и случайном состоянии равном 42.

Метод голосования

На конечном этапе выходы парциальных классификаторов подаются на блок жесткого голосования (англ. Hard voting) [15].

Жесткое голосование — это метод голосования в машинном обучении, при котором выбирается класс с наибольшим большинством голосов.

Жесткое голосование объединяет выходные данные различных алгоритмов и назначает наиболее часто встречающийся результат в качестве окончательного.

Пусть $O_1,O_2,O_3,....,O_n$ — выходы различных моделей для обнаружения аномалий. Алгоритм жесткого голосования выбирает тот выход, который встречается чаще всего среди всех результатов используемых моделей. Механизм жесткого голосования объединяет их результаты, синтезируя более надежное и достоверное представление кластеризации. Это позволяет создать более точную интерпретацию базовой структуры данных, усиливая точность модели и её способность обнаруживать аномалии. Поскольку в данной реализации задействованы три модели - следовательно, итоговый результат будет определяться по принципу "двукратного появления" одного из двух классов — аномалии или нормального состояния, что обеспечивает бинарную классификацию данных.

Метрики анализа производительности

Для сбора результатов экспериментов используются следующие стандартные метрики оценки производительности. В следующем списке метрик оценки производительности TP - истинно положительные, FP - ложно положительные, TN - истинно отрицательные, FN - ложно отрицательные.

- Точность (англ. Accuracy) рассчитывается как доля правильно классифицированных объектов — как положительных, так и отрицательных

— по отношению ко всем предсказаниям, включая ошибочные. Формула для расчета точности выглядит следующим образом:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Точность (англ. Precision) определяется как отношение количества правильно классифицированных объектов, принадлежащих определённому классу, к общему количеству объектов, которые модель предсказала как принадлежащие этому классу.

$$Precision = \frac{TP}{TP + FP}$$

- Recall — это отношение количества правильно идентифицированных положительных изображений к общему количеству положительных изображений, включая ложноотрицательные.

$$Recall = \frac{TP}{TP + FN}$$

- F-мера (англ. F1-Score) — это среднее гармоническое значение точности и отзыва на основе предсказаний, сделанных моделью.

$$F1\text{-}Score = 2 * \frac{Precission * Recall}{Precission + Recall}$$

Заключение

В данной работе предложена система обнаружения вторжений, основанная на неконтролируемом обучении, для выявления DDoS-атак нулевого дня в системах IoT. В качестве классификатора был использован ансамбль из трех алгоритмов кластеризации: метод k-средних, стохастический градиентный спуск одно-классового метода опорных векторов и модель гауссовской смеси. Этот ансамбль был применен для классификации экземпляров данных как нормальных или атакованных. Предложенная модель была оценена с использованием стандартных метрик

производительности, таких как точность, полнота, точность положительных предсказаний и F-мера. Модель показала точность 94,50 % и F-мера 94,30 %, что свидетельствует о минимальном числе ложноположительных и ложноотрицательных результатов [7]. В будущем подобный способ обнаружения атак нулевого дня может быть расширен как по числу кластеризаторов входящих в ансамбль (рис.1), так и для обнаружения других типов кибератак нулевого дня, таких как веб-атаки или атаки грубой силы.

Литература

1. Рыбаков, С. Ю. Методы защиты IoT от атак нулевого дня // Инженерный вестник Дона. – 2025. – № 3. URL: ivdon.ru/ru/magazine/archive/n3y2025/9944

2. Li Q., Huang H., Li R., Lv J., Yuan Z., Ma L., Han Y., Jiang Y. A comprehensive survey on DDoS defense systems: New trends and challenges. Computer Networks, 2023, pp. 1-25.

3. Bilge L., Dumitra T. Before we knew it: an empirical study of zero-day attacks in the real world // In Proceedings of the 2012 ACM conference on Computer and communications security, 2012, pp. 833–844.

4. Parkinson S., Khan S. Identifying irregularities in security event logs through an object-based chi-squared test of independence // Journal of information security and applications, 2018, vol. 40, pp. 52–62.

5. Zoppi T., Ceccarelli A., Salani L., Bondavalli A. On the educated selection of unsupervised algorithms via attacks and anomaly classes // Journal of Information Security and Applications, 2020, vol. 52, pp. 1-11.

6. Aldweesh A., Derhab A., Emam A.Z. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 2020, vol. 189, pp. 105-124.

7. Vu K., Poirion P.-L., Liberti L. Gaussian random projections for Euclidean membership problems // Discrete Applied Mathematics, 2019, vol. 253, pp. 93–102.

8. Roopak M., et al. An unsupervised approach for the detection of zero-day distributed denial of service attacks in Internet of Things networks // IET Netw. 2024, pp. 1–15.

9. Hartigan J. A., Wong M. A. Algorithm as 136: A k-means clustering algorithm // Journal of the royal statistical society. series c (applied statistics), 1979, vol. 28, no. 1, pp. 100–108.

10. Blanco R., Malagón P., Briongos S., Moya J. M. Anomaly detection using gaussian mixture probability model to implement intrusion detection system // In Hybrid Artificial Intelligent Systems: 14th International Conference, HAIS 2019, León, Spain, September 4–6, 2019, Proceedings 14. Springer, 2019, pp. 648–659.

11. Hu H., Liu J., Zhang X., Fang M. An effective and adaptable k-means algorithm for big data cluster analysis. Pattern Recognition, 2023, vol. 139, pp. 1-18.

12. Manevitz L.M., Yousef M. One-class svms for document classification // Journal of machine Learning research, 2001, vol. 2, no. Dec, pp. 139–154.

13. Guo Y. A review of machine learning-based zero-day attack detection: Challenges and future directions // Computer Communications, 2022, pp. 175-185.

14. Huang Z., Gou Z. Gaussian mixture model-based pattern recognition for understanding the long-term impact of covid-19 on energy consumption of public buildings // Journal of Building Engineering, 2023, vol. 72, pp. 1-20.

15. Dai Q.-Y., Zhang B., Dong S.-Q., et al. A ddos-attack detection method oriented to the blockchain network layer // Security and Communication Networks, 2022, pp. 1-18.

References

1. Rybakov, S. Y. Inzhenernyi vestnik Dona, 2025. № 3. URL: ivdon.ru/ru/magazine/archive/n3y2025/9944

2. Li Q., Huang H., Li R., Lv J., Yuan Z., Ma L., Han Y., Jiang Y. A comprehensive survey on DDoS defense systems: New trends and challenges. Computer Networks, 2023, pp. 1-25.

3. Bilge L., Dumitra T. Proceedings of the 2012 ACM conference on Computer and communications security, 2012, pp. 833–844.

4. Parkinson S., Khan S. Journal of information security and applications, 2018, vol. 40, pp. 52–62.

5. Zoppi T., Ceccarelli A., Salani L., Bondavalli A. / Journal of Information Security and Applications, 2020, vol. 52, pp. 1-11.

6. Aldweesh A., Derhab A., Emam A.Z. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 2020, vol. 189, pp. 105-124.

7. Vu K., Poirion P.-L., Liberti L. Discrete Applied Mathematics, 2019, vol. 253, pp. 93–102.

8. Roopak M., et al. IET Netw. 2024, pp. 1–15.

9. Hartigan J. A., Wong M. A. Journal of the royal statistical society. series c (applied statistics), 1979, vol. 28, no. 1, pp. 100–108.

10. Blanco R., Malagón P., Briongos S., Moya J. M. Hybrid Artificial Intelligent Systems: 14th International Conference, HAIS 2019, León, Spain, September 4–6, 2019, Proceedings 14. Springer, 2019, pp. 648–659.

11. Hu H., Liu J., Zhang X., Fang M. An effective and adaptable k-means algorithm for big data cluster analysis. Pattern Recognition, 2023, vol. 139, pp. 1-18.

12. Manevitz L.M., Yousef M. Journal of machine Learning research, 2001, vol. 2, no. Dec, pp. 139–154.

13. Guo Y. Computer Communications, 2022, pp. 175-185.

14. Huang Z., Gou Z. Journal of Building Engineering, 2023, vol. 72, pp. 1-20.

15. Dai Q.-Y., Zhang B., Dong S.-Q., et al. Security and Communication Networks, 2022, pp. 1-18.

Дата поступления: 17.03.2025

Дата публикации: 25.06.2025