Научная статья
DOI 10.66424/2071-8217-2026-2-4
УДК 004.056

ОБНАРУЖЕНИЕ АНОМАЛИЙ В СОБЫТИЯХ БЕЗОПАСНОСТИ ОС НА ОСНОВЕ СТАТИСТИЧЕСКОГО АНАЛИЗА И БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

А. В. Немчинов[1], Т. Д. Овасапян[2*], Е. В. Жуковский[2]

[1]Санкт-Петербургский государственный университет телекоммуникаций им. профессора М. А. Бонч-Бруевича, Санкт-Петербург, Россия

[2]Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

✉ *otd@ibks.spbstu.ru*

ДЛЯ ЦИТИРОВАНИЯ

Немчинов А. В., Овасапян Т. Д., Жуковский Е. В. Обнаружение аномалий в событиях безопасности ОС на основе статистического анализа и больших языковых моделей // Проблемы информационной безопасности. Компьютерные системы. 2026. № 2. С. 49–59.
DOI: 10.66424/2071-8217-2026-2-4

ПОСТУПИЛА 27.04.2026
ПРИНЯТА 06.05.2026
ОПУБЛИКОВАНА 15.06.2026

АННОТАЦИЯ

Исследованы возможности применения больших языковых моделей и статистического анализа для автоматизации обнаружения аномалий в событиях безопасности ОС. Предложен метод обнаружения аномалий, позволяющий автоматически выделять значимые отклонения и формировать их интерпретацию. Разработан программный прототип, реализующий данный метод, и проведено его тестирование.

КЛЮЧЕВЫЕ СЛОВА

Статистический анализ, большие языковые модели, обнаружение аномалий

Original article
DOI 10.66424/2071-8217-2026-2-4

DETECTING ANOMALIES IN SECURITY EVENTS BASED ON STATISTICAL ANALYSIS AND LARGE LANGUAGE MODELS

A. V. Nemchinov[1], T. D. Ovasapyan[2*], E. V. Zhukovsky[2]

[1]St. Petersburg State University of Telecommunication Named After Professor M. A. Bonch-Bruevich, St. Petersburg, Russia

[2]Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

✉ *otd@ibks.spbstu.ru*

FOR CITATION

Nemchinov A. V., Ovasapyan T. D., Zhukovsky E. V. Detecting anomalies

ABSTRACT

The possibilities of using large language models and statistical methods to automate the detection of anomalies in OS security events

in security events based on statistical analysis and large language models. *Problems of information security. Computer systems*. 2026. No. 2, pp. 49–59. DOI: 10.66424/2071-8217-2026-2-4 (In Russian)

RECEIVED 27.04.2026
ACCEPTED 06.05.2026
PUBLICATION 15.06.2026

are investigated. A method for detecting anomalies is proposed that allows to automatically identify significant deviations and form their interpretation. A software prototype implementing this method has been developed and tested.

KEYWORDS

Statistical analysis, large language models, anomaly detection

1. ВВЕДЕНИЕ

Каждый компонент инфраструктуры генерирует события безопасности: журналы доступа, сообщения от средств защиты, системные логи. Security Information and Event Management (SIEM) – системы сбора и корреляции событий безопасности ежедневно обрабатывают миллионы таких событий. Каждое из них потенциально может свидетельствовать об инциденте безопасности. Однако большинство событий являются следствием легитимной работы системы или ложными срабатываниями [1].

Ручной анализ таких больших объемов данных требует значительных трудозатрат и высокой квалификации специалиста. Даже при использовании стандартных правил корреляции оператор получает множество предупреждений. Возникает потребность в автоматизированных средствах, которые могли бы не только выделять подозрительные отклонения в потоке событий, но и представлять их в удобной для восприятия форме, снижая нагрузку на специалиста [1, 2].

Активно развиваются методы обнаружения аномалий на основе статистического анализа [3, 4]. Они применяются для отсеивания фоновой активности и выделения статистически значимых отклонений. Параллельно с этим большие языковые модели демонстрируют хорошие результаты в понимании и генерации текста [5]. Их применение открывает возможности для автоматической интерпретации выявленных отклонений и формирования понятных оператору сообщений. Однако прямое использование LLM для анализа всего потока событий нецелесообразно из-за высоких вычислительных затрат и риска потери значимых сигналов в шуме [1].

Таким образом, актуальной задачей является разработка комбинированного подхода, сочетающего статистическую фильтрацию для выделения ограниченного числа значимых аномалий и последующую их обработку с помощью LLM. Это позволит снизить нагрузку на оператора, предоставляя ему только действительно важные отклонения с понятными пояснениями.

2. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБНАРУЖЕНИЯ АНОМАЛИЙ

Задача обнаружения аномалий заключается в выявлении ключевых данных, событий или наблюдений, которые значительно отклоняются от нормального набора данных [6, 7]. В контексте анализа событий безопасности под аномалией понимается резкое изменение интенсивности событий определенного типа или нарушение статистических закономерностей.

Поскольку события безопасности регистрируются во времени, их удобно представлять в виде временных рядов – последовательностей значений, упорядоченных в порядке времени. Далее будут рассмотрены основные из применяемых статистических подходов, основанных на анализе этих временных рядов.

Пороговые методы. Самым простым является установление фиксированных порогов. Если значение некоторой метрики, например, количество событий в минуту, превышает заданный порог T, фиксируется аномалия. Пороги могут быть

относительными или абсолютными. Абсолютные пороги задаются в виде фиксированных числовых значений и не зависят от контекста функционирования системы. Относительные пороги, наоборот, вычисляются на основе нормального поведения системы, например, как отклонение от среднего значения за определенный период.

Достоинством метода является простота реализации и минимальные вычислительные затраты. Однако фиксированные пороги не адаптируются к изменениям в поведении системы, требуют ручной настройки и могут приводить к большому числу ложных срабатываний.

Методы скользящего среднего (SMA и EWMA). Более гибкими являются методы, использующие скользящие статистики [8, 9]. Их суть в построении сглаженной оценки текущего уровня ряда на основе недавних наблюдений. Эта оценка будет прогнозом ожидаемого значения, а отклонение фактического наблюдения от прогноза может свидетельствовать об аномалии. Простое скользящее среднее (SMA) вычисляется как среднее арифметическое последних k наблюдений:

$$\bar{x}_t = \frac{1}{k} \sum_{i=t-k+1}^{t} x_i,$$

где x_i – количество событий в i-й период времени; x_t – количество событий за фиксированный период времени t; k – количество последовательных наблюдений.

Аномалия формируется при $\left|x_t - \bar{x}_{t-}\right| > \delta$. Параметр δ выбирается эмпирически, например, как удвоенное стандартное отклонение значений внутри окна.

Экспоненциально взвешенное скользящее среднее (EWMA) придает больший вес свежим данным:

$$\hat{x}_t = ax_t + (1-a)\hat{x}_{t-1},$$

где $a \in (0,1)$ – параметр сглаживания. Чем ближе a к единице, тем быстрее модель реагирует на новые наблюдения.

Отклонение оценивается по правилу:

$$\left|x_t - \hat{x}_{t-1}\right| > \beta\sigma,$$

где σ – оценка стандартного отклонения процесса; β – коэффициент.

Если неравенство выполняется, наблюдение помечается как аномальное. Метод EWMA хорошо реагирует на резкие изменения, однако не учитывает сезонные колебания – регулярные, повторяющиеся изменения метрики во времени. Например, в выходные и праздничные дни интенсивность событий снижается, тогда как в рабочие дни она выше. В результате такие закономерные изменения могут ошибочно интерпретироваться как аномалии.

Метод кумулятивных сумм (CUSUM). Метод кумулятивных сумм предназначен для обнаружения небольших постоянных смещений среднего [10]. Накопленные отклонения вверх и вниз рассчитываются по формулам:

$$S_t^+ = \max\left(0, S_{t-1}^+ + x_t - \mu_0 - k\right),$$

$$S_t^- = \max\left(0, S_{t-1}^- - x_t + \mu_0 - k\right),$$

где μ_0 – целевое среднее (ожидаемое значение при отсутствии нарушений); k – допустимое отклонение.

Аномалия фиксируется при превышении порога h $\left(S_t^+ > h,\ S_t^- > h\right)$. CUSUM эффективен для обнаружения длительного повышения количества событий, однако чувствителен к сезонным колебаниям, если они не учтены при выборе μ_0.

Робастные статистики. Робастные статистики устойчивы к наличию выбросов, в отличие от рассматриваемых [7, 11]. Наиболее распространенные робастные статистики – медиана и медианное абсолютное отклонение (MAD). Медиана M – центральное значение упорядоченной выборки, половина значений меньше, половина больше. MAD определяется как медиана абсолютных отклонений наблюдений от медианы:

$$M = median\left\{x_1, x_{2,\ldots}, x_n\right\},$$

$$\mathrm{MAD} = median\left\{\left|x_i - M\right|\right\}.$$

На основе медианы и MAD строится модифицированная Z-оценка. Она означает число стандартных отклонений, на которое значение наблюдаемой величины отклоняется от медианы [6]:

$$Z = \frac{|x_t - M|}{1{,}4826\,\mathrm{MAD}},$$

где x_t – количество событий за фиксированный период времени t.

Коэффициент 1,4826 возникает из отношения между MAD и стандартным отклонением для нормального распределения [12]. Для нормально распределенной случайной величины справедливо равенство:

$$\mathrm{MAD} \approx 0{,}6745\sigma.$$

Обратное соотношение позволяет выразить стандартное отклонение через MAD:

$$\sigma \approx \frac{\mathrm{MAD}}{0{,}6745} \approx 1{,}4826\,\mathrm{MAD}.$$

Таким образом, коэффициент 1,4826 приводит MAD к масштабу стандартного отклонения в предположении нормального распределения. Это позволяет интерпретировать Z так же, как классическую Z-оценку. Значения $Z > 3$ обычно считаются признаком аномалии.

Достоинства данного подхода заключаются в устойчивости к единичным выбросам в данных, отсутствии предположений о распределении, в простоте вычисления и наглядности получаемых результатов.

Каждый метод имеет свои сильные и слабые стороны. Для использования выбран метод робастных статистик и Z-оценки, так как имеет наибольшую устойчивость к выбросам при сохранении интерпретируемости и низкой вычислительной сложности.

3. ОПИСАНИЕ ПРЕДЛАГАЕМОГО МЕТОДА ОБНАРУЖЕНИЯ АНОМАЛИЙ

Исходными данными для работы метода будут события безопасности, поступающие от различных источников инфраструктуры. В SIEM-системе они агрегируются по двум ключевым признакам: источнику событий и типу событий. Для каждой такой пары ведется учет количества срабатываний за фиксированные временные интервалы. Получаемые временные ряды сохраняются и будут использоваться для построения профиля нормального поведения [9]. Общая схема этого процесса показана на рис. 1.

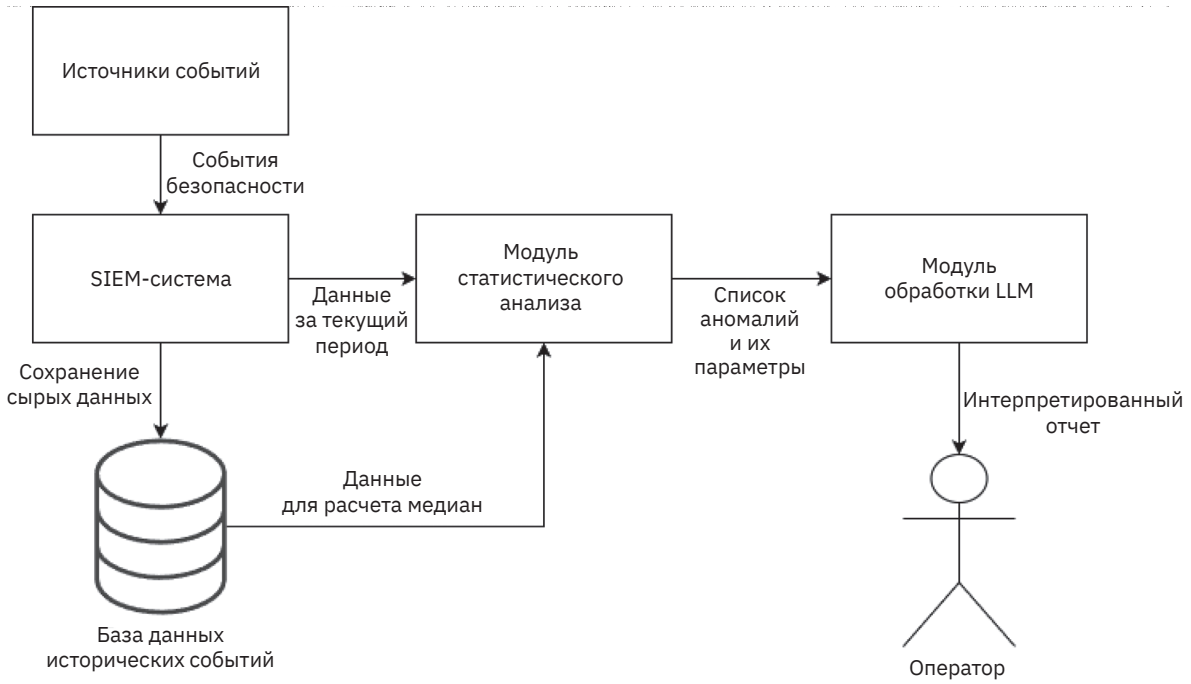

Рис. 1 | Общая схема работы метода обнаружения аномалий

Fig. 1 | General scheme of operation of the anomaly detection method

В качестве временного интервала выбраны сутки – такой промежуток будет самым оптимальным для анализа из-за периодичности активности в инфраструктуре. Для учета сезонных колебаний все дни разделяются на две категории: рабочие и выходные, так как характер и количество событий в эти дни существенно отличается [3, 9].

Для каждого источника и типа событий отдельно для рабочих и выходных дней вычисляются медиана и медианное абсолютное отклонение. Медиана M отражает типичное количество событий данного типа в сутки, а MAD характеризует естественный разброс вокруг медианны. Эти показатели станут стандартом, на который будут ссылаться показания в новый день.

При обработке очередных суточных данных для каждого текущего значения x_t его отклонение от «стандартных» значений. На основе ранее вычисленных переменных вычисляется модифицированная Z-оценка, она показывает, насколько удалено текущее значение от типичного. Однако при малых значениях медианны даже небольшое абсолютное изменение может давать высокую Z-оценку, что приведет к ложным срабатываниям. Чтобы избежать этого, решение о том является ли наблюдение аномалией принимается на основе нескольких критериев.

Первый критерий – абсолютное отклонение $\Delta = |x_t - M|$. Оно должно превышать некоторое минимальное значение $\Delta_{\min}$ – это отсеивает случайные отклонения от медианны.

Второй критерий – относительное отклонение в процентах $\delta = \frac{\Delta}{M} \cdot 100\,\%$, при $M > 0$. Если медиана равна нулю, то δ принимается равным 100 % при наличии событий. Также применяется порог $\delta_{\min}$, он зависит от величины медианны. Для малых медиан порог устанавливается выше, около 150 %, для больших и средних примерно 40 и 80 % соответственно.

Третий критерий – модифицированная Z-оценка, для которой тоже устанавливается порог $Z_{\min}$. При малых медианах $Z_{\min} = 3$, при средних – 4, при больших – 5. Это обеспечивает адаптивность к разным масштабам данных и позволяет избежать ложных срабатываний из-за естественной активности.

Наблюдение фиксируется как аномалия только при выполнении перечисленных условий. Количество аномалий может быть велико даже при таком отсеивании статистического шума. Поэтому аномалии также ранжируются по уровню критичности, который определяется весовым коэффициентом. Он учитывает величину Z-оценки, относительное отклонение и уровень критичности событий, присвоенной SIEM-системой. Например, правила с уровнем критичности 10 и выше получают повышающий множитель. На рис. 2 приведен алгоритм статистического анализа выявления аномалий.

Отобранные аномалии структурируются и передаются большой языковой модели. Языковая модель способна оценить семантический контекст происходящего [1, 2] и анализирует совокупность передаваемых параметров: идентификатор источника, тип события, текущее количество, медиану, величину отклонения и критичность события. На их основе языковая модель формирует список аномалий, которые с наибольшей вероятностью связаны с инцидентами безопасности.

4. РЕАЛИЗАЦИЯ ПРОГРАММНОГО КОМПЛЕКСА И ОЦЕНКА ЭФФЕКТИВНОСТИ

Разработанный программный комплекс реализован на языке Python и интегрирован с SIEM-системой Wazuh. Он состоит из трех модулей, взаимодействующих через файловую систему и API Telegram.

1. **Сбор и агрегация.** Ежедневно запускается по расписанию. Выполняет аутентификацию в Wazuh Indexer и выгружает события за последние сутки с помощью scroll-запросов. Это позволяет обрабатывать большие объемы событий.

Все сырые события сохраняются в сжатом формате gzip в каталоге raw_events,

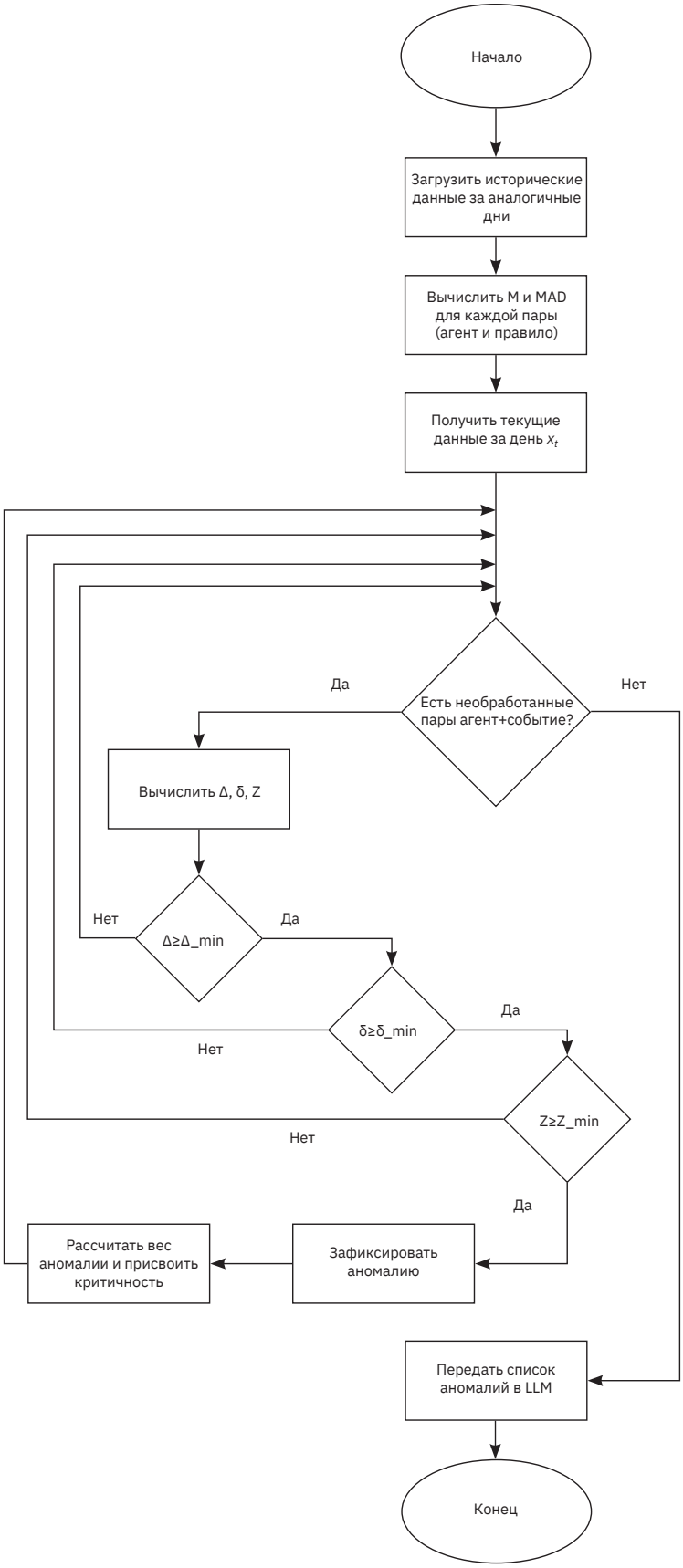

Рис. 2 | Алгоритм статистического выявления аномалий

Fig. 2 | An algorithm for statistical detection of anomalies

что дает возможность при необходимости пересчитать статистику или изменить алгоритм агрегации без повторного обращения к Wazuh Indexer. Также формируется агрегированная статистика для каждого агента и правила, подсчитывается количество срабатываний. Результат записывается в json-файл в каталоге stats_events.

2. **Обнаружение аномалий.** Реализуется метод обнаружения аномалий. На вход принимает статистику за прошедшие сутки. Сначала определяется тип дня – рабочий или выходной. Затем выбираются аналогичные предшествующие дни, количество определяется по параметру HISTORY_DAYS, по умолчанию пять дней. Для каждого дня загружается соответствующий статистический файл. Для каждой пары «агент – правило» формируется временной ряд: по выбранным предыдущим дням собираются значения количества срабатываний соответствующего правила. На основе этого ряда вычисляются медиана M и медианное абсолютное отклонение MAD. Если срабатывание правила у агента встречалось менее чем в двух днях, то MAD принимается равным нулю.

Далее загружается статистика текущего дня и последовательно проверяются все записи. Для каждой записи вычисляются абсолютное отклонение $\Delta = |x_t - M|$, относительное отклонение $\delta = \frac{\Delta}{M} \cdot 100\ \%$, модифицированная Z-оценка $Z = \frac{|x_t - M|}{1{,}4826\,\mathrm{MAD}}$. Запись признается аномалией, если одновременно все перечисленные параметры проходят пороги (табл. 1).

Каждой аномалии присваивается вес:

$$w = Z + \frac{\delta}{100} f,$$

где f – коэффициент важности правила; Z – модифицированная Z-оценка; δ –относительное отклонение.

Коэффициент важности определяется так: 2 для уровня правила ≥ 10; 1,5 для уровня 7–9; 1 для уровня меньше 7. Это позволяет выделить наиболее важные отклонения одновременно по уровню опасности события и по величине отклонения. Аномалии сортируются по убыванию веса, и в итоговый список попадают только самые весомые.

Все пороговые значения задаются через переменные окружения, это позволяет гибко настраивать алгоритм под конкретную инфраструктуру. Полный список настраиваемых параметров приведен в табл. 2.

Результатом работы второго модуля является json-файл today_stat. Он содержит дату, общее число найденных аномалий и детальную информацию про каждую. Этот файл сохраняется и одновременно копируется в отдельную директорию для передачи третьему модулю.

Третий модуль – интеграция с LLM и оповещение. Формируется подробное сообщение с найденными аномалиями для оператора на основании файла today_stat. Аномалиям присваивается уровень опасности исходя из рассчитанного веса: критический при $w \geq 20$, высокий при $10 \leq w < 20$, средний при $5 \leq w < 10$ и низкие при $w < 5$. Для каждой аномалии пишется имя агента, его ip и id в Wazuh, номер правила, уровень опасности,

Таблица 1 | Адаптивные пороги в зависимости от медианы

Table 1 | Adaptive thresholds depending on the median

Диапазон медиан	Порог абсолютного отклонения Δ_{min}	Порог относительного отклонения δ_{min},%	Порог Z-оценки z_{min}
$M < 20$	10	150	3,0
$20 \leq M < 50$		80	4,0
$50 \leq M < 200$		40	5,0
$M \geq 200$		Не используется	6,0

Таблица 2 | Основные параметры конфигурации модуля обнаружения аномалий

Table 2 | The main configuration parameters of the anomaly detection module

Переменная	Значение по умолчанию	Описание
HISTORY_DAYS	5	Количество похожих дней
MIN_ABSOLUTE_DIFF	10	Минимальная абсолютное отклонение
MIN_PERCENT_DIFF_SMALL	150	Порог относительного отклонения при медиане <20
MIN_PERCENT_DIFF_SMALL_PLUS	80	Порог относительного отклонения при медиане 20–50
MIN_PERCENT_DIFF_MEDIUM	40	Порог относительного отклонения при медиане 50–200
Z_SCORE_THRESHOLD_SMALL	3,0	Порог *Z* при медиане <20
Z_SCORE_THRESHOLD_SMALL_PLUS	4,0	Порог *Z* при медиане 20–50
Z_SCORE_THRESHOLD_MEDIUM	5,0	Порог *Z* при медиане 50–200
Z_SCORE_THRESHOLD_LARGE	6,0	Порог *Z* при медиане >200
MAX_ANOMALIES_PER_AGENT	5	Максимум аномалий для агента
MAX_TOTAL_ANOMALIES	50	Общее число аномалий в сообщении

текущее количество событий, вес и относительное отклонение. Сформированное сообщение отправляется оператору.

Далее содержимое файла today_stat передается в большую языковую модель DeepSeek с использованием API. Для наиболее точного ответа используется промпт, который предписывает модели выбирать аномалии на основе числовых показателей и семантики событий, переводить описание событий на русский язык и формировать краткое сообщение. Это сообщение также отправляется оператору, как краткая выжимка наиболее опасных аномалий.

Помимо модуля обнаружения аномалий реализован модуль ежедневной статистики. Он формирует общую сводку по состоянию SIEM-системы. В сводку включаются количество активных и отключившихся агентов, топ наиболее часто встречающихся событий за сутки, число исправленных и новых уязвимостей на основе данных сканера уязвимостей Wazuh. На основе данных за последние 10 дней строятся два графика, показывающих изменение количества событий и уязвимостей, а также их распределение по уровням критичности. Пример графиков можно увидеть на рис. 3.

Тестирование разработанного комплекса проводилось в реальной инфраструктуре, включающей 60 активных агентов SIEM. Ежедневно в Wazuh регистрировалось около миллиона событий. Модуль статистического обнаружения аномалий выявлял в среднем 14 аномалий в сутки. После обработки результатов большой языковой моделью сообщение для оператора содержало информацию о 2–3 наиболее значимых аномалиях. Все события, которые впоследствии были квалифицированы как инциденты безопасности, вошли в число отобранных LLM. Это свидетельствует об эффективности данного подхода и разработанного программного комплекса.

5. ЗАКЛЮЧЕНИЕ

В результате выполнения работы исследованы методы статистического обна-

ружения аномалий в событиях безопасности, проведен их анализ. Обоснован выбор робастных статистик в сочетании с модифицированной *Z*-оценкой как наиболее устойчивого к выбросам и интерпретируемого подхода. Разработан метод обнаружения аномалий, учитывающий сезонные колебания активности и использующий адаптивные пороги. Предложен весовой коэффициент, позволяющий ранжировать аномалии с учетом всех числовых величин.

На основе предложенного метода реализован программный комплекс, интегрированный с Wazuh и Telegram. Комплекс состоит из трех модулей: сбора и агрегации событий, обнаружения аномалий с помощью статистических методов,

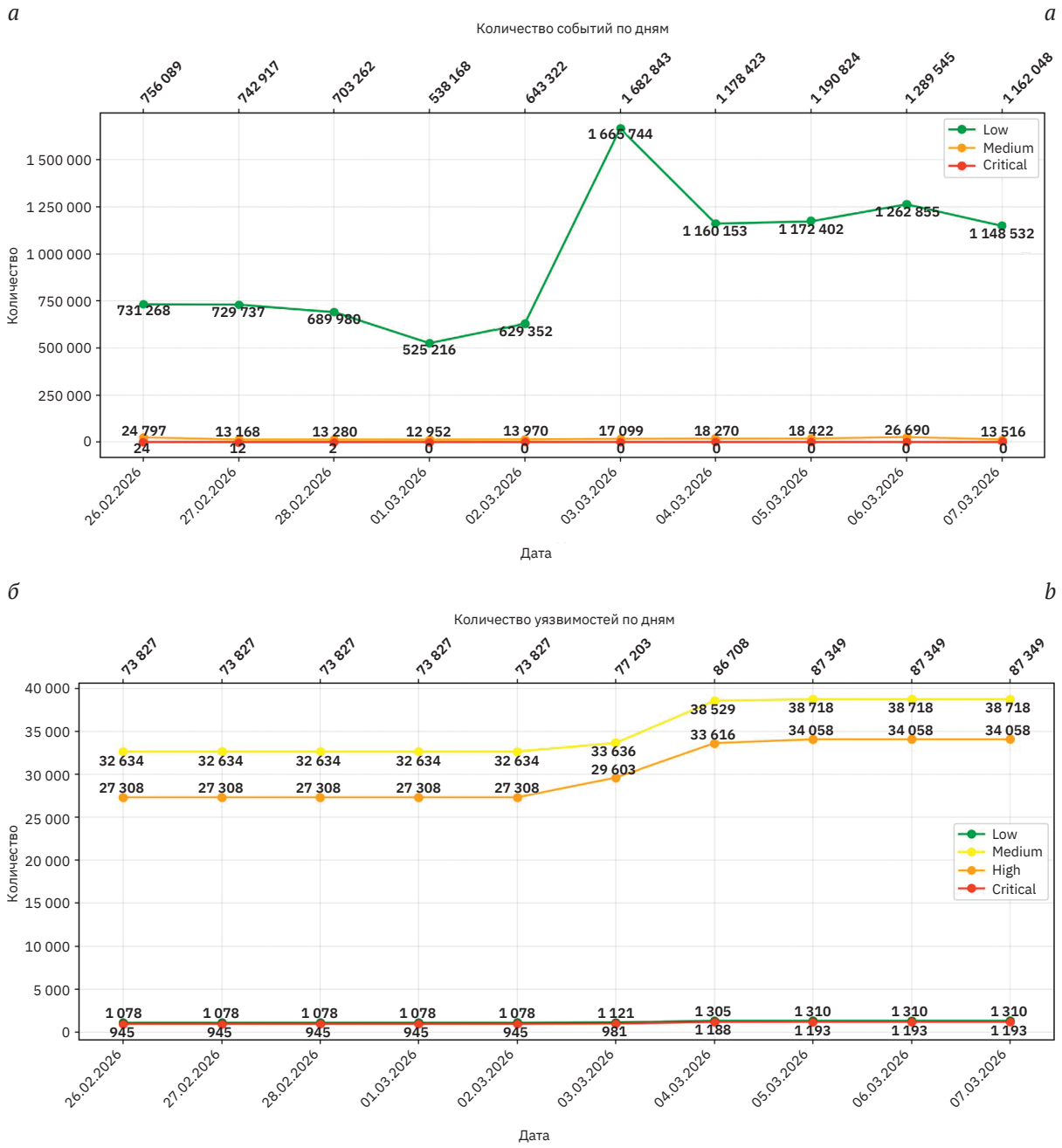

Рис. 3 | График распределения событий (*а*) и уязвимостей (*б*) по 10 последним дням

Fig. 3 | Schedule of distribution of events (*a*) and vulnerabilities (*b*) for the last 10 days

а также модуля интеграции с большой языковой моделью DeepSeek для фильтрации и интерпретации выявленных аномалий. Разработан дополнительный модуль ежедневной статистики, помогающий отслеживать активность в SIEM, предоставляя оператору контекстную информацию.

Экспериментальная проверка проведена в условиях реальной эксплуатации и подтвердила работоспособность и эффективность предложенного подхода. Применение LLM позволило значительно сократить число оповещений и расширить понимание аномального поведения. При этом сохраняется полнота выявленных инцидентов. Дальнейшее развитие комплекса предполагает автоматическое обновление эталонных значений медиан и MAD по мере накопления новых данных, учет многомерных корреляций между событиями, а также расширение функциональности языковой модели для написания рекомендаций по реагированию на выявленные аномалии.

КОНФЛИКТ ИНТЕРЕСОВ / CONFLICT OF INTERESTS

Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interests.

СПИСОК ИСТОЧНИКОВ

1. **Chagna Y., Goldschmidt A.** Next-generation cyberattack detection with large language models: anomaly analysis across heterogeneous logs // arXiv preprint arXiv:2602.06777. 2026.
2. **da Silva G. de J. C., Westphall C. B.** A survey of large language models in cybersecurity // arXiv preprint arXiv:2402.16968. 2024.
3. **Hochenbaum J., Vallis O. S., Kejariwal A.** Automatic anomaly detection in the cloud via statistical learning // arXiv preprint arXiv: 1704.07706. 2017.
4. **Stetsyuk M., Anikin V., Pyrch O. et al.** Method of detecting anomalies in IoT device traffic based on statistical analysis using the modified Z score // CEUR Workshop Proceedings. 2025. Vol. 3963. P. 284–298.
5. **Minaee S., Mikolov T., Nikzad N. et al.** Large language models: A Survey // arXiv preprint arXiv:2402.06196. 2024.
6. **Стельмах Н. Е., Козачков А. В.** Обзор методов выявления аномалий при аудите системных вызовов в ОС // International Journal of Open Information Technologies. 2025. Т. 13. № 9. С. 25–33.
7. **Iglewicz B., Hoaglin D. C.** How to detect and handle outliers. Milwaukee: ASQC Quality Press, 1993. 87 p.
8. **Roberts S. W.** Control chart tests based on geometric moving averages // Technometrics. 1959. Vol. 1. № 3. P. 239–250.
9. **Vallis O. S., Hochenbaum J., Kejariwal A.** A novel technique for long-term anomaly detection in the cloud // 6th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 14). 2014.
10. **Page E. S.** Continuous inspection schemes // Biometrika. 1954. Vol. 41. № 1–2. P. 100–115.
11. **Leys C., Ley C., Klein O. et al.** Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median // Journal of Experimental Social Psychology. 2013. Vol. 49. № 4. P. 764–766.
12. **Rousseeuw P. J., Croux C.** Alternatives to the median absolute deviation // Journal of the American Statistical Association. 1993. Vol. 88. № 424. P. 1273–1283.

REFERENCES

1. **Chagna Y., Goldschmidt A.** Next-generation cyberattack detection with large language models: anomaly analysis across heterogeneous logs. *arXiv preprint arXiv:2602.06777*. 2026.

2. **da Silva G. de J. C., Westphall C. B.** A survey of large language models in cybersecurity. *arXiv preprint arXiv:2402.16968*. 2024.
3. **Hochenbaum J., Vallis O. S., Kejariwal A.** Automatic anomaly detection in the cloud via statistical learning. *arXiv preprint arXiv: 1704.07706*. 2017.
4. **Stetsyuk M., Anikin V., Pyrch O. et al.** Method of detecting anomalies in IoT device traffic based on statistical analysis using the modified Z score. *CEUR Workshop Proceedings*. 2025. Vol. 3963, pp. 284–298.
5. **Minaee S., Mikolov T., Nikzad N. et al.** Large language models: A Survey. *arXiv preprint arXiv:2402.06196*. 2024.
6. **Stelmach N. E., Kozachok A. V.** Review of anomaly detection methods during system call auditing in OS. *International Journal of Open Information Technologies*. 2025. Vol. 13. No. 9, pp. 25–33. (In Russian)
7. **Iglewicz B., Hoaglin D. C.** How to detect and handle outliers. Milwaukee: ASQC Quality Press, 1993, 87 p.
8. **Roberts S. W.** Control chart tests based on geometric moving averages. *Technometrics*. 1959. Vol. 1. No. 3, pp. 239–250.
9. **Vallis O. S., Hochenbaum J., Kejariwal A.** A novel technique for long-term anomaly detection in the cloud. 6th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 14). 2014.
10. **Page E. S.** Continuous inspection schemes. *Biometrika*. 1954. Vol. 41. No. 1–2, pp. 100–115.
11. **Leys C., Ley C., Klein O. et al.** Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*. 2013. Vol. 49. No. 4, pp. 764–766.
12. **Rousseeuw P. J., Croux C.** Alternatives to the median absolute deviation. *Journal of the American Statistical Association*. 1993. Vol. 88. No. 424, pp. 1273–1283.

СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

НЕМЧИНОВ Александр Владимирович – студент, Санкт-Петербургский государственный университет телекоммуникаций им. профессора М. А. Бонч-Бруевича, Россия, 193232, Санкт-Петербург, пр-т Большевиков, д. 22, к. 1
E-mail: sasha01082004@gmail.com
ORCID: 0009-0001-3348-5038

NEMCHINOV Alexander V. – Student, St. Petersburg State University of Telecommunication Named After Professor M. A. Bonch-Bruevich, Russia, 193232, St. Petersburg, Bolshevikov ave., 22, bldg. 1

ОВАСАПЯН Тигран Джаникович – канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29
E-mail: otd@ibks.spbstu.ru
ORCID: 0000-0002-2009-5460

OVASAPYAN Tigran D. – Candidate of Engineering Sciences, Associate Professor, Peter the Great St. Petersburg Polytechnic University, Russia, 195251, St. Petersburg, Polytechnicheskaya str., 29

ЖУКОВСКИЙ Евгений Владимирович – канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29
E-mail: ezhukovsky@ibks.spbstu.ru
ORCID: 0009-0006-9013-1750

ZHUKOVSKY Evgeny Vladamirovich. – Candidate of Engineering Sciences, Associate Professor, Peter the Great St. Petersburg Polytechnic University, Russia, 195251, St. Petersburg, Polytechnicheskaya str., 29