УДК 004.056
Шифр специальности: ВАК 2.3.6

А.В. Иванов, М.А. Киселев

Методика оценки качества логирования событий информационной безопасности в итерационно-управленческом цикле с применением метрики дефицита логирования

Актуальность. Качество логирования событий информационной безопасности определяет полноту последующей нормализации, корреляции и анализа инцидентов в SIEM-системах. Неполнота или структурная некорректность событий снижает наблюдаемость инфраструктуры и затрудняет реагирование на инциденты. ***Цель исследования.*** Разработать методику количественной оценки качества логирования событий информационной безопасности на основе интегрального показателя дефицита логирования *DL* и итерационного цикла PDCA. ***Методы.*** Использованы формализация профиля логирования, матрица требований к регистрации событий, набор обязательных полей, метрики полноты и точности логирования, рекуррентная модель динамики *DL* и представление методики в нотации IDEF0. ***Научная новизна.*** Предложена методика, отличающаяся совместным использованием матрицы требований к логированию, набора обязательных полей, интегрального показателя *DL* и межитерационного учёта новых и устранённых дефицитов в цикле PDCA. ***Результаты.*** На примере потока событий аутентификации Windows показано снижение *DL* с 0,30 до 0,11 за пять итераций при росте числа корректных записей с 700 до 890. ***Практическая значимость.*** Методика может применяться для формализованного контроля качества логирования, приоритезации корректирующих мероприятий и включения оценки качества событий в регулярные процессы мониторинга ИБ.

Ключевые слова: PDCA, SIEM, оценка качества логирования событий ИБ, полнота и точность логирования, дефицит логирования, математическая модель, итеративные методологии, ISO 27001, NIST CSF.
DOI: 10.21293/1818-0442-2026-29-1-114-123

Введение

Функционирование современных систем информационной безопасности основывается на качественном логировании событий, к которому предъявляются повышенные требования к качеству логирования событий, выступающие основой для дальнейшего анализа инцидентов информационной безопасности, последующей нормализации и корреляции в системах класса SIEM (Security Information and Event Management). Неполное, нерегулярное логирование нарушает целостность информационной картины происходящих процессов, затрудняет реализацию механизмов автоматического обнаружения аномалий и значительно снижает эффективность реагирования на инциденты [1, 2].

По данным отраслевых исследований и отчетов по инцидентам, значительная часть атак остаётся без должного расследования из-за отсутствия необходимых событий ИБ. Согласно отчету IBM «Cost of a Data Breach Report 2023», среднее время на обнаружение и локализацию инцидента (жизненный цикл утечки) составило 277 дней [3].

Недостаточная видимость и неполнота событий ИБ являются одними из ключевых факторов, увеличивающих это время. Это подтверждается аналитикой Cisco Talos, согласно которой в I квартале 2024 г. одной из главных проблем при реагировании на инциденты стало отсутствие необходимых событий, что осложнило расследования в 28,6% случаев [4]. В рамках настоящей работы предполагается, что наблюдаемый дефицит событий обусловлен сочетанием технических и организационных факторов: отсутствием формализованных требований к профилю логирования, нестабильностью / несогласованностью доставки событий и отсутствием регулярной процедуры контроля качества логов.

Анализ существующих подходов [4, 6] показал, что управление логированием чаще носит реактивный характер и опирается на частные проверки качества событий ИБ. В руководствах по управлению качеством событий ИБ, как правило, оценивается по набору разрозненных частных признаков: полноте охвата источников и классов событий, наличию (поступлению) ожидаемых записей за заданный период, а в меньшей степени – корректности (валидности) формата и содержания записей. Дополнительно акцент делается на контроле того, что все активы, включённые в контур мониторинга, действительно генерируют и передают события ИБ в систему сбора/анализа (NIST SP 800-92; ENISA) [7, 8]. В связи с этим в настоящей работе вводится интегральный показатель *DL* дефицита логирования, отражающий одновременно пропуски событий и ошибки структуры записей и предназначенный для количественной оценки динамики качества логирования событий ИБ.

В профессиональной и научной литературе представлены близкие по назначению подходы к оценке полноты и качества журналов событий. Наиболее близким является направление log completeness, ориентированное на оценку наличия ожидаемых записей в журнале [9]. Кроме того, рассматриваются показатели семантической корректности и согласованности записей логов [10], а в нормативных и методических источниках фиксируются требования к управлению журналами безопасности и составу регистрируемой информации [8, 11]. Вместе

с тем для задач контроля качества логирования в SIEM-контуре недостаточно учитывать только факт поступления ожидаемых событий: необходимо также оценивать структурную валидность записей, наличие обязательных полей и пригодность событий для последующей нормализации, корреляции и анализа. Поэтому в настоящей работе показатель *DL* рассматривается как интегральная мера остаточного несоответствия фактического состояния логирования заданному профилю требований.

Термины и определения

Событие информационной безопасности (далее – событие ИБ): зафиксированное в обрабатываемом виде состояние информационной (автоматизированной) системы, сетевого, телекоммуникационного, коммуникационного, иного прикладного сервиса или информационно-телекоммуникационной сети, указывающее на возможное нарушение целостности, доступности и (или) конфиденциальности информации, а также на сбой в работе средства защиты/обработки информации или иную ситуацию, которая может быть значимой для безопасности информации [11].

Под *полнотой логирования C* в рамках настоящего исследования понимается степень охвата всех логируемых событий, подлежащих регистрации в соответствии с внутренними политиками организации, архитектурными требованиями и международными стандартами, включая NIST SP 800-92 [8] и ISO/IEC 27001:2022 [5]. К таким событиям относятся, в частности, попытки аутентификации, операции доступа к объектам, действия по изменению или удалению ресурсов, сетевые соединения и изменения конфигурации.

Точность логирования A в рамках данного исследования определяется как соответствие записей фактическим событиям по содержанию, структуре и метаинформации. Согласно NIST SP 800-92 [8], корректная запись должна содержать временную метку (UTC), идентификаторы субъекта и объекта действия, тип операции, результат выполнения, источник события (например, IP-адрес) и тип журнала.

Качество логирования – интегральная характеристика, отражающая степень соответствия фактически регистрируемых событий ИБ профилю логирования организации (матрица требований **R** и набор обязательных полей *P**), сформированному на основе требований ГОСТ Р 59548–2022 [11], рекомендаций (например, NIST SP 800-92 [8]) и внутренних регламентов мониторинга ИБ.

Класс события (событий ИБ) – это категория однородных событий, объединённых общим смыслом действия/нарушения и потому имеющих единые требования к регистрации: ожидаемую частоту/объём за период *T* и набор обязательных полей для проверки полноты и валидности записи.

*Набор обязательных полей P** – базовый набор атрибутов записи, используемый для проверки валидности события по соответствующей паре «класс источника/класс события»; *P** может расширяться в зависимости от задач мониторинга и особенностей источников событий ИБ.

Матрица требований **R** задаёт соответствие «источник × класс событий» и определяет перечень классов, обязательных к регистрации для каждого контролируемого источника в рамках профиля логирования (с учётом нормативных требований и внутренних регламентов). В практической постановке элемент $\mathbf{R}[j, s; T]$ трактуется как ожидаемое число записей класса *j* от источника *s* за период *T*.

Дефицит логирования (*DL*, Deficit of Logging) – интегральный показатель отклонения фактического состояния логирования от эталонного профиля **R**, *P** за период *T*, учитывающий одновременно пропуски событий и нарушения валидности записей; *DL* применяется для количественной оценки динамики качества логирования в итерационно-управленческом цикле.

Профиль логирования – формализованное описание требований к регистрации событий ИБ для контролируемой инфраструктуры, включающее матрицу требований **R**, набор обязательных полей *P**, период анализа *T*, а также пороговые значения полноты, точности и допустимого дефицита логирования.

Период анализа T – фиксированный временной интервал, за который выполняются сбор, нормализация, валидация событий ИБ и расчёт показателей качества логирования.

Ожидаемое число записей N – ожидаемое количество записей контролируемых классов событий, которое должно быть получено от активных источников за период *T* в соответствии с матрицей требований **R**.

Наблюдаемое число записей N_{obs} – количество фактически полученных за период *T* записей, относящихся к контролируемым классам событий по матрице **R**, после предварительной нормализации и удаления дубликатов, но до проверки обязательных полей *P**.

Реестр дефицитов – перечень выявленных несоответствий профилю логирования, включающий источник события, класс события, тип нарушения, признак новизны или устранения нарушения и связь с корректирующим мероприятием.

Фактическое число корректных записей $C_{\text{факт}}$ – количество записей из числа наблюдаемых событий, которые удовлетворяют требованиям набора обязательных полей *P**, имеют валидные значения ключевых атрибутов и корректную временную метку в пределах периода анализа *T*.

Эталонное число записей $C_{\text{эталон}}$ – ожидаемое число записей контролируемых классов событий за период *T*, задаваемое матрицей требований **R**. В практических расчётах далее обозначается как *N*.

Новые отклонения n_i – количество новых нарушений полноты или валидности логирования, выявленных на i-й итерации.

Устранённые отклонения n_i – количество нарушений логирования, устранённых на *i-й* итерации по результатам корректирующих мероприятий.

Пороговые значения $C_{\min}, A_{\min}, DL_{\max}$ – минимально допустимые значения полноты и точности

логирования, а также максимально допустимое значение дефицита логирования, используемые для принятия решения о продолжении или завершении корректирующих мероприятий.

Постановка проблемы

Отсутствие критически важных записей/полей в событиях ИБ приводит к снижению качества логирования событий ИБ [10].

Целью настоящего исследования является разработка и апробация методики количественной оценки качества логирования событий информационной безопасности, основанной на интеграции показателя *DL* и его использовании в итерационном управленческом цикле.

Практический эффект применения методики состоит в формализованном контроле полноты и валидности логируемых событий и в возможности количественно оценивать остаточный дефицит логирования и его изменение от итерации к итерации. Результаты расчёта используются для приоритезации корректирующих мероприятий и включения контроля качества логирования в регулярные процессы мониторинга ИБ.

Научная новизна работы заключается в разработке формализованной методики оценки качества логирования событий ИБ, отличающейся совместным использованием матрицы требований к логированию **R**, набора обязательных полей *P**, интегрального показателя дефицита логирования *DL* и межитерационного учёта новых и устранённых дефицитов в цикле PDCA. В нормативных и методических источниках основное внимание уделяется требованиям к регистрации событий, составу регистрируемой информации, охвату источников и обеспечению поступления событий в систему сбора и анализа [8, 11], а в отдельных исследованиях рассматриваются показатели полноты и семантической корректности логов [9, 10]. Отличие предлагаемой методики состоит в том, что *DL* рассматривается не как самостоятельный аналог log completeness, а как интегральный показатель остаточного дефицита логирования, связывающий матрицу требований **R**, набор обязательных полей *P**, период анализа *T*, метрики полноты и валидности записей, а также межитерационную динамику устранения нарушений в цикле PDCA.

Количественная модель динамики дефицита логирования в процессе автоматизации оценки качества логирования событий ИБ

В данной части работы разрабатывается количественная модель, описывающая динамику изменения показателя дефицита логирования во времени при итеративном применении PDCA-цикла [12].

Цикл PDCA был выбран на основе следующих причин:

- PDCA обеспечивает итерационную организацию работ по улучшению качества логирования и позволяет формализовать процесс через последовательность Plan–Do–Check–Act.
- Параметры предложенной модели *DL* имеют прямое численное соответствие этапам PDCA (новые дефициты, устранённые дефициты, остаточный *DL*), что обеспечивает воспроизводимую количественную оценку эффекта каждой итерации и динамики изменений во времени.
- Использование PDCA упрощает включение контроля качества логирования в регулярные процессы мониторинга ИБ и последующую автоматизацию расчётов и отчётности в SIEM.

Модель позволяет фиксировать остаточный дефицит логирования событий ИБ, количественно оценивать эффект от реализованных мероприятий и планировать корректирующие действия на последующие итерации.

Предложенная модель позволяет описывать динамику изменения показателя *DL* в процессе последовательного выполнения итераций PDCA. *DL* определяется как относительное несоответствие текущего состояния логирования установленным требованиям по полноте, точности логирования.

Формализация понятия дефицита логирования

Показатель *DL* агрегирует нарушения двух типов:

1)дефицит полноты;

2)дефицит точности.

Значения *DL* лежат в диапазоне [0; 1]: при $DL = 0$ дефицит отсутствует, при $DL \to 1$ наблюдается критическая неполнота и/или искажение логируемых событий.

Для формализации профиля логирования используются два значения:

- матрица требований **R**;
- набор обязательных полей *P**.

В рамках статьи *P** рассматривается на уровне укрупнённых групп атрибутов (в терминах нормализованной схемы события):

- метка времени в UTC и идентификатор записи / события;
- идентификация источника (узел / сервис / актив);
- идентификация субъекта (учётная запись / процесс/сессия);
- идентификация объекта воздействия (ресурс / узел-получатель / объект БД и т.п.);
- действие и результат (включая код/ошибку при наличии);

При этом обязательность конкретных полей фиксируется локальным профилем требований к логированию, который задаёт обязательные поля отдельно для разных типов источников и классов событий.

Формула базового дефицита логирования

В наиболее простой постановке *DL* определяется как 1 минус доля корректных записей от их эталонного количества за период *T*:

$$DL = 1 - \left(C_{\text{факт}} / C_{\text{эталон}} \right). \quad (1)$$

Источники, временно отключённые от мониторинга или не введённые в **R**, в расчёт $C_{\text{эталон}}$ и $C_{\text{факт}}$ за данный период не включаются.

Величина $C_{\text{эталон}}$ получается суммированием заданных в **R** значений по всем контролируемым классам и активным источникам.

$$N = C_{\text{эталон}}(T) = \sum_{j,s} \mathbf{R}[j, s; T],$$

$$DL = \max\left(0, 1 - C_{\text{факт}} / C_{\text{эталон}}\right).$$

Формирование модели динамики *DL* в зависимости от характеристик системы логирования

Формирование модели динамики *DL* определяется особенностями инфраструктуры, свойствами поступающих событий ИБ и требованиями к чувствительности оценки качества логирования событий ИБ. Таким образом, возможны два варианта формирования модели динамики *DL*.

• *Нормированная линейная модель* применяется в условиях устойчивой, предсказуемой среды, где задержки в доставке событий минимальны, а корректирующие мероприятия дают быстрый и отчётливый эффект. Такая модель наиболее подходяща при регулярной доставке событий ИБ, стабильной нагрузке и наличии оперативного доступа к источникам событий. Она обеспечивает высокую чувствительность к изменениям и позволяет быстро отслеживать прогресс итераций PDCA, особенно в тестовых или контролируемых средах.

• *Модель с экспоненциальным сглаживанием* предпочтительна в условиях неустойчивой среды, когда устранение дефицита логирования происходит с задержкой (например, из-за рассинхронизации времени, буферизации событий, неравномерного сбора или агрегации событий). Эта модель снижает влияние временных аномалий, не отражающих устойчивую тенденцию, фильтрует шум и позволяет более реалистично учитывать инерционность изменений.

Под нормированной моделью понимаем рекуррентное обновление *DL* в долях от $C_{\text{эталон}}$; $(n_i - c_i)/N$ делает метрику безразмерной [0;1] и сопоставимой между источниками и периодами. Нормированная рекуррентная модель

$$DL_i = DL_{i-1} - (c_i - n_i)/N. \quad (2)$$

Ограничения для DL_i:

$$DL_i = \max\left\{0, \min\left\{1, DL_{i-1} - (c_i - n_i)/N\right\}\right\}.$$

Базовая нормированная линейная модель динамики дефицита логирования *DL* предполагает, что изменения состояния происходят мгновенно и линейно: каждое улучшение (устранение ошибок) немедленно снижает значение показателя *DL*. Однако на практике устранение дефицита логирования может быть распределено во времени, зависеть от задержек доставки событий ИБ, интервалов сбора и других факторов.

Под «шумом» в настоящей модели понимаются кратковременные нерегулярные отклонения в объёме поступающих событий, не связанные с реальными изменениями качества логирования, а вызванные техническими причинами: задержками доставки, буферизацией, ротацией логов, несинхронизацией времени, временными сбоями каналов передачи и т. п. Такие отклонения могут существенно влиять на метрику *DL* при использовании линейной модели и, как следствие, искажают эффект итерации. Для сглаживания подобных искажений применяется модель с экспоненциальным сглаживанием, в которой коэффициент α подбирается в зависимости от уровня таких колебаний в среде логирования.

В таких случаях полезно использовать модифицированную модель с экспоненциальным сглаживанием.

Определения (нормированная постановка):

• DL_i – значение дефицита логирования на текущей итерации;

• $\alpha \in [0;1]$ – коэффициент инерционности модели. Он определяет, насколько сильно новая итерация влияет на обновлённое значение DL_i.

Чем выше α, тем более «инерционно» ведёт себя модель: медленно реагирует на изменения, но хорошо сглаживает шум.

Чем ниже α, тем быстрее модель реагирует на колебания, но становится чувствительной к единичным всплескам и артефактам.

• $\Delta_i^{(N)} = c_i - n_i$ – чистый эффект итерации (положителен, если устранений больше, чем новых дефицитов логирования). Нормированная постановка:

$$\Delta_i^{(N)} = (c_i - n_i)/N. \quad (3)$$

Для удобства восприятия при построении сглаженной модели используется обозначение $\Delta_i^{(N)}$, соответствующее чистому положительному эффекту итерации: разности между устранёнными c_i и вновь выявленными n_i дефицитами логирования. Сглаженная модель задаётся выражением

$$DL_i = \alpha DL_{i-1} + (1-\alpha)\left(DL_{i-1} - \Delta_i^{(N)}\right), \quad (4)$$

что эквивалентно:

$$DL_i = DL_{i-1} - (1-\alpha)(c_i - n_i)/N. \quad (5)$$

Из (5) следует, что скорость уменьшения *DL* пропорциональна балансу устранённых и новых дефицитов логирования, а степень реакции регулируется параметром α.

Рекомендации по выбору α. Для сред, где возможны задержки при передаче и обработке событий (например, при удалённой отправке логов по протоколу syslog или через буферизующие конвейеры типа Logstash), целесообразно выбирать α в диапазоне 0,7–0,95 с целью сглаживания кратковременных всплесков. Напротив, при локальной, практически мгновенной записи событий в журналы (например, с использованием службы аудита Linux auditd или локального домена journald без промежуточной агрегации) имеет смысл задать меньшее значение α (около 0,2–0,5) для более быстрой реакции модели на изменения.

Методика оценки полноты и точности логирования и их интеграции в показатель *DL*

Методика предназначена для воспроизводимой оценки качества логирования событий информационной безопасности за фиксированный период T и

последующего включения результатов в цикл улучшений методологии PDCA. Интегральным показателем выступает *DL*, определяемый формулой (1). Для практического расчёта *DL* вводятся метрики полноты и точности логирования и регламентируется порядок их вычисления.

Входные данные:

- матрица требований к логированию **R**;
- набор обязательных полей *P**;
- события за период *T*, приведённые к единому формату, без дубликатов и с нормализованным временем.

Формируется **R** по результатам анализа текущей инфраструктуры и лог-активности:

– по типам источников (Windows, Linux, БД, СЗИ);

– по количеству таких источников (из CMDB или реестра SIEM);

– по нормативно установленной минимально допустимой частоте появления событий каждого класса за период *T*, определяемой на основе статистики прошлых периодов и требований политики безопасности.

При этом для каждого класса событий выбирается пороговое значение, которое отражает нижнюю границу нормальной активности (например, если в норме фиксируется 50–100 событий в час, порог устанавливается не ниже 50).

Формирование матрицы **R** выполняется на этапе Plan цикла PDCA и может рассматриваться как отдельная процедура подготовки профиля логирования. На практике **R** реализуется как конфигурационный справочник требований к логированию: для каждой группы источников и класса событий задаётся селектор (например, EventID) и эталонное условие ожидаемости на период *T* (минимальная частота/порог). При изменении инфраструктуры (добавление/удаление узлов, новых типов журналов) матрица **R** должна пересматриваться и актуализироваться, чтобы эталонное количество событий отражало реальное состояние контролируемой среды.

Выходные данные:

- сквозные метрики за *T*: полнота $C_{\text{полнота}}$, точность *A*, интегральный показатель *DL*;
- реестр дефицитов (несоответствий) с межитерационными идентификаторами n_i (новые) и c_i (устранённые) нарушения;
- решения по фазам PDCA, пороговые значения и критерии остановки.

Определения метрик полноты и точности: пусть *N* – ожидаемое число событий за период *T* (эталон по матрице **R**).

Введём: $C_{\text{полнота}} = N_{\text{obs}}/N,\ 0 \le C_{\text{полнота}} \le 1$,

- полнота логирования:

$$C_{\text{полнота}} = N_{\text{obs}}/N,\ 0 \le C_{\text{полнота}} \le 1. \qquad (6)$$

На данном этапе не выполняется проверка на соответствие правилам *P** – в частности, обязательные поля могут быть неполными или содержать невалидные значения. Это позволяет использовать N_{obs} в качестве общей совокупности, внутри которой вычисляется доля корректных записей;

- точность логирования:

$$A = C_{\text{факт}}/N_{\text{obs}},\ 0 \le A \le 1, \qquad (7)$$

где запись считается корректной, если удовлетворяет правилам *P** (обязательные поля заполнены, значения валидны, временная метка в границах *T* и в UTC).

Тогда $C_{\text{факт}} = A \cdot N_{\text{obs}}$ и (1) эквивалентна интерпретационному виду

$$DL = 1 - C_{\text{факт}}/N = 1 - A \cdot C_{\text{полнота}}. \qquad (8)$$

Соотношение (8) показывает связь предлагаемого показателя *DL* с известным понятием *log completeness*. Компонент $C_{\text{полнота}}$ отражает полноту поступления ожидаемых событий за период T и соответствует той части оценки, которая близка к *log completeness*. Компонент *A* отражает долю валидных записей среди наблюдаемых событий и учитывает структурную корректность записей относительно набора обязательных полей *P**. При $A = 1$ показатель *DL* принимает вид $1 - C_{\text{полнота}}$ и вырождается в дефицит полноты поступления событий. При $C_{\text{полнота}} = 1$ показатель *DL* принимает вид $1 - A$ и отражает дефицит валидности записей. В общем случае *DL* является интегральной оценкой совместного влияния двух типов нарушений: непоступления ожидаемых событий и структурной непригодности поступивших записей для последующей обработки в SIEM. Поэтому *DL* не заменяет показатель *log completeness*, а использует его содержательную основу как один из компонентов более общей модели качества логирования.

Дополнительное отличие *DL* состоит в привязке оценки к профилю логирования конкретной инфраструктуры: матрице требований **R**, набору обязательных полей *P**, периоду анализа *T* и пороговым значениям $C_{\min}, A_{\min}, DL_{\max}$. В отличие от частных метрик полноты или семантической корректности логов, *DL* применяется в итерационном цикле PDCA и позволяет количественно отслеживать динамику устранения дефицитов за счёт учёта новых и устранённых нарушений между итерациями.

Для помодульного учёта по классам событий *j* допускается агрегация по классам событий (равновесовое усреднение)

$$DL = 1 - \frac{1}{m}\sum_{j=1}^{m} A_j C_{\text{полнота},j}, \qquad (9)$$

где *m* – это количество контролируемых классов событий в матрице **R**.

Порядок расчёта за период *T*:

1. Утвердить **R**, *P**, пороги. $C_{\min}, A_{\min}, DL_{\max}$, $C_{\min}$ – минимально допустимая полнота логирования за период *T*: доля покрытых требуемых событий $C_{\text{полнота}} = N_{\text{obs}}/N$ должна удовлетворять $C_{\text{полнота}} \ge C_{\min}$; $A_{\min}$ – минимально допустимая точность логирования: доля валидных записей среди

наблюдённых: $A = C_{\text{факт}} / N_{\text{obs}}$ должна удовлетворять $A \geq A_{\min}$.

$DL_{\max}$ – максимально допустимый интегральный дефицит логирования: по (1) $DL = 1 - C_{\text{факт}} / N$ должна удовлетворять $DL \leq DL_{\max}$, что эквивалентно

$$C_{\text{факт}} / N \geq 1 - DL_{\max}.$$

Пороговые условия согласованности. Для обеспечения корректности вычисления: $C_{\text{полнота}} = \min\left(N_{\text{obs}} / N, 1\right)$ при $N_{\text{obs}} = 0$ полагаем $A = 0$; классы с $N = 0$ в расчет периода T не включаются.

Пороговые значения *DL* согласуются напрямую с формулой $DL = 1 - A \cdot C_{\text{полнота}}$. При заданном целевом $DL_{\max}$ пары $C_{\min}, A_{\min}$ должны удовлетворять

$$C_{\min} \cdot A_{\min} \geq 1 - DL_{\max}.$$

Симметричное решение

$$C_{\min} = A_{\min} = \sqrt{1 - DL_{\max}}.$$

Если один порог задан заранее, второй определяется дробью

$$C_{\min} = (1 - DL_{\max}) / A_{\min}.$$

2. Собрать события за период *T*; удалить дубликаты; перевести метки времени в UTC; привести поля к единой схеме (*P**); классифицировать записи по классам из матрицы **R**. Результат – нормализованный набор за *T*, размеченный по классам.

3. По матрице **R** задать эталонное количество записей *N* за период *T* (суммируя значения по классам/источникам. Подсчитать N_{obs} – фактически полученные за *T* записи, относящиеся к классам **R**, до проверки обязательных полей.

4. Применить правила *P**; число прошедших проверку – $C_{\text{факт}}$ и $C_{\text{факт},\,j}$.

5. Вычислить метрики:\
$C_{\text{полнота}} = N_{\text{obs}} / N, A = C_{\text{факт}} / N_{\text{obs}}, DL = 1 - C_{\text{факт}} / N$.

6. Сравнить с порогами $C_{\min}, A_{\min}, DL_{\max}$ и сформировать корректирующие мероприятия.

Определение целевых значений для $C_{\min}, A_{\min}, DL_{\max}$:

1. Фиксация желаемого уровня $DL_{\max} - 0{,}05$ (не более 5% интегральных потерь/искажений за период *T*).

2. Из тождества $DL = 1 - A \cdot C_{\text{полнота}}$ следует условие согласованности порогов

$$C_{\min} \cdot A_{\min} \geq 1 - DL_{\max} = 0{,}95.$$

$C_{\min} = A_{\min} = \sqrt{0{,}95} \approx 0{,}975$. В рамках данной статьи используется допустимая асимметрия порогов: $C_{\min} = 0{,}95$, $A_{\min} = 0{,}98$. Поскольку $C_{\min} \cdot A_{\min} = 0{,}931 < 0{,}95$, критерий $DL \leq 0{,}05$ применяется как самостоятельное ограничение наряду с условиями $A \geq 0{,}98$ и $C_{\text{полнота}} \geq 0{,}95$.

Итерация признаётся успешной, если одновременно выполняются три неравенства:

$$C_{\text{полнота}} \geq C_{\min}, A \geq A_{\min}, DL \leq DL_{\max}.$$

Представление результатов и критерии остановки. По завершении расчётов за период *T* формируется единый пакет отчётности (формирование отчётности и критериев завершения относится к этапу представления результатов).

1. Карта метрик по классам событий (включённым в матрицу **R**). Для каждого класса *j* приводятся: N_i – ожидаемое число записей; $C_{\text{факт}}$ – наблюдённое число; $N_{\text{obs},j}$ – число валидных записей по *P**.

2. Метрики:

$$C_{\min}, A_{\min}, DL_{\max}, C_{\text{полнота},\,j}, A_j, DL_j.$$

3. Реестр дефицитов: c_i, n_i.

Фаза насыщения и критерии остановки цикла PDCA

После нескольких итераций вклад очередных исправлений может становиться минимальным: остаётся остаточный *DL*, обусловленный техническими артефактами сбора и доставки (задержки, буферизация и повторные отправки, ротация логов, рассинхронизация времени, ошибки дедупликации/парсинга, кратковременные сбои каналов), а также ограничениями платформы и допустимыми исключениями, закреплёнными в политике логирования. Это и есть фаза насыщения.

Признаки насыщения:

1. $|DL_i - DL_{i-1}| < \varepsilon$ на протяжении *k* последовательных итераций (типичное $\varepsilon = 0{,}01 - 0{,}02$ и $k = 2 - 3$).

2. Оставшиеся нарушения локализованы в немногих классах/источниках и требуют несоразмерных усилий (например, вендорские ограничения).

Практический критерий завершения при насыщении: если все ключевые классы выполняют пороги, а агрегированное $DL \leq DL_{\max}$ и $|DL_i - DL_{i-1}| < \varepsilon$ в течение *k* итераций, цикл завершается; оставшиеся точечные дефициты переводятся в эксплуатационный реестр корректирующих работ.

Для формализации последовательности действий методика представлена в нотации IDEF0. На контекстной диаграмме A0 (рис. 1) рассматривается обобщённая функция «Оценить качество логирования событий информационной безопасности за период *T*». Входами функции являются реестр контролируемых источников и активов, а также события ИБ за период *T*. Управляющими воздействиями выступают нормативные и внутренние требования, пороговые значения показателей $C_{\min}, A_{\min}, DL_{\max}$, а также правила PDCA и критерии остановки. Механизмами выполнения являются SIEM/ETL-контур, средства анализа и расчёта метрик, эксперт SOC или администратор. Выходами функции являются рассчитанные *C*, *A*, *DL* за период *T*, реестр выявленных дефицитов и причин, а также корректирующие меры и решение PDCA. Декомпозиция функции A0 (рис. 2) включает четыре этапа: формирование профиля логирования; сбор и нормализацию событий; расчёт показателей *C*,

A, *DL* и выявление дефицитов; формирование корректирующих мер и решения PDCA.

На этапе A1 формируется профиль логирования, включающий матрицу требований **R**, набор обязательных полей *P** и период анализа *T*. На этапе A2 события ИБ за период *T* собираются и приводятся к нормализованному виду. На этапе A3 рассчитываются показатели полноты, точности и дефицита логирования, а также формируется реестр выявленных дефицитов. На этапе A4 по результатам расчёта формируются корректирующие меры и решение о продолжении или завершении итерации PDCA. Выход O3 используется не только как результат текущей итерации, но и как обратная связь для уточнения профиля логирования и корректирующих действий на следующей итерации.

Пример расчёта динамики *DL*

Период наблюдения $T = 1$ ч. Сбор выполняется через Windows Event Forwarding на коллектор wec-app1. Рассматривается класс событий «Аутентификация», включающий события из журнала Security: 4624, 4625, 4634, 4672, 4768, 4769, 4776.

Задание эталона: **R** и расчёт *N*: пусть $\mathbf{H} = \{\mathbf{h_1}, \ldots, \mathbf{h_{10}}\}$ – множество контролируемых узлов домена. Для каждого узла нормативно задана минимальная активность не менее 100 аутентификаций в час (успешных и неуспешных), тогда:

$$\mathbf{R}[\text{auth}, \text{h}_k, T] = 100, \quad k = 1, \ldots, 10,$$

$$N = \sum_{k=1}^{10} \mathbf{R}[\text{auth}, \text{h}_k, T] = 10 \cdot 100 = 1000.$$

Здесь *N* – эталонное число записей по матрице **R** за период *T*.

Задание правил валидности *P**: для класса «Аутентификация» словарь *P** задаётся в терминах нормализованной схемы события и включает следующий минимум:

$$P^* = \{\text{timestamp, msgid, subject.id,}$$
$$\text{account, domain, dst.asset,}$$
$$\text{logon\_type, logon\_service,}$$
$$\text{status, src/dst}\}.$$

Пороговые значения для PDCA: $C_{\min} = 0,95$; $A_{\min} = 0,98$, $DL_{\max} = 0,05$.

Базовое измерение (до исправлений): по итогам сбора и нормализации получено $N_{\text{obs}} = 880$ записей требуемого класса (дефицит покрытия 120). Валидация по *P** выявила 180 некорректных/неполных записей, следовательно, $C_{\text{факт}} = 700$.

$$C_{\text{полнота}} = 0,88, A = 0,79, DL = 0,30.$$

Типовые причины дефицита логирования: неполный охват WEF (в том числе Kerberos 4768/4769), выключенный аудит Logon на части хостов, несоответствие времени (локальная TZ вместо UTC), неполная нормализация 4624 (не извлечены LogonType/AuthenticationPackage), отсутствие разрешения активов (только IP без hostname).

Корректирующие мероприятия (фаза Do):

- WEF/подписки: включить WinRM, обновить подписку Security по EventID {4624, 4625, 4634, 4672, 4768, 4769, 4776}; исключить фильтрацию, приводящую к потере LogonType = 3 (Network).

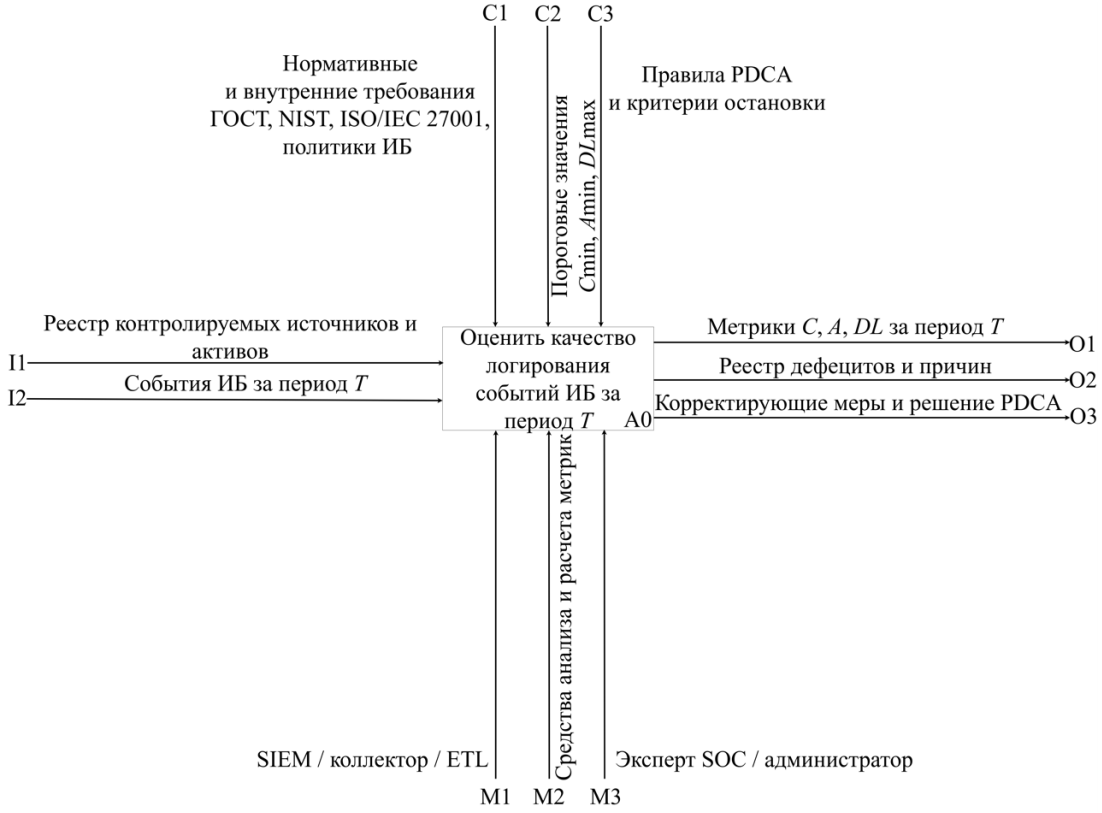

Рис. 1. Контекстная диаграмма A0 методики оценки качества логирования событий ИБ за период *T* в нотации IDEF0

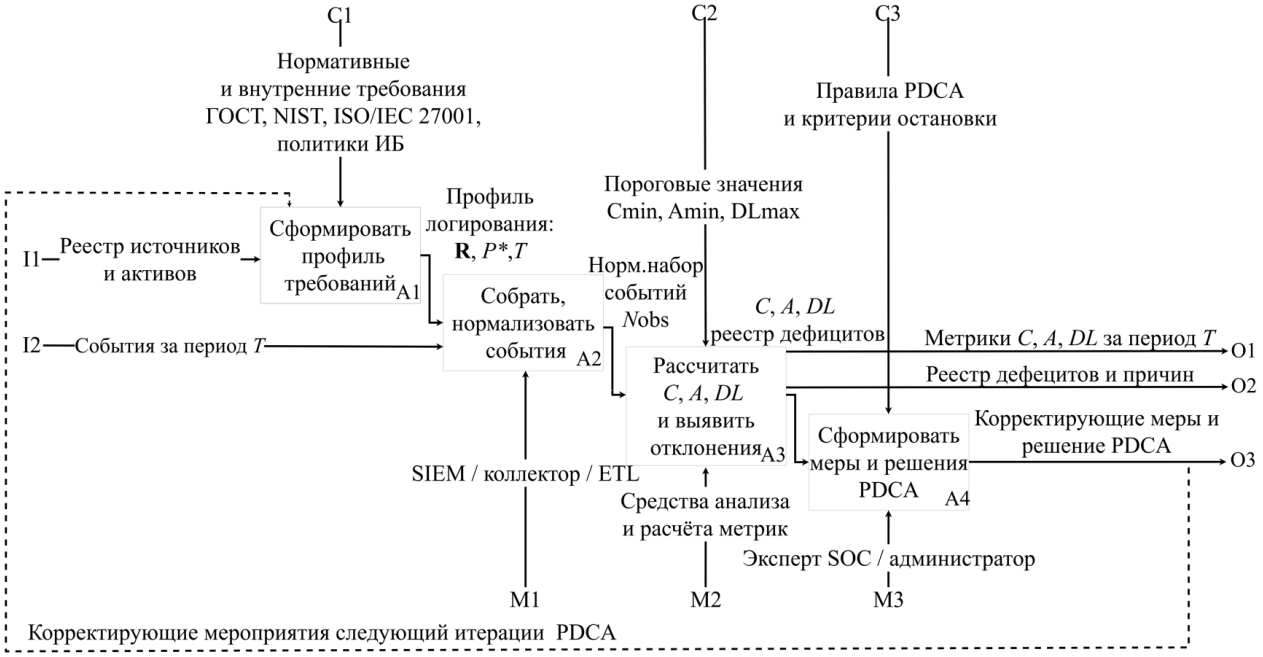

Рис. 2. Декомпозиция функции A0 «Оценить качество логирования событий ИБ за период *T*» в нотации IDEF0

• Аудит на хостах: включить Success/Failure для «Logon/Logoff» и верифицировать настройку.

• Нормализация/ETL: приводить время к UTC; разбирать поля 4624 (LogonType→logon_type, AuthenticationPackage→logon_service); классифицировать учётные записи с «$» как machine; обогащать src/dst asset по инвентарю и DNS (IP→FQDN→asset).

Динамика изменений на фазе Check оценивалась с использованием рекуррентной модели при фиксированном эталонном числе записей. Результаты расчёта показателя *DL* по итерациям PDCA приведены в таблице.

$$DL_i = DL_{i-1} + (n_i - c_i)/N, \quad N = 1000.$$

Динамика DL_i по итерациям PDCA

№ *i*	Меры изменений	n_i	c_i	DL_i	$C_{\text{факт}}$
0	Исходное состояние	–	–	0,300	700
1	Обновлены WEF-подписки (4624/4768/4769), включён аудит Logon	10	50	0,260	740
2	Нормализация 4624 (LogonType, Kerberos), UTC-время	12	55	0,217	783
3	Обогащение src/dst asset; классификация *$ как machine	9	48	0,178	822
4	Подключены ранее неохваченные узлы; скорректированы правила отбора событий (фильтрация) в подписках	7	42	0,143	857
5	Точечные правки (W32Time, DNS-записи)	5	38	0,110	890

Интерпретация: за пять итераций *DL* снизился с 0,3 до 0,11, рост ($C_{\text{факт}}: 700 \to 890$). Порог $DL_{\max} = 0{,}05$ пока не достигнут; цикл PDCA продолжается.

Решение (фаза Act) и критерии завершения: итерация считается успешной при одновременном выполнении $C_{\text{полнота}} \geq C_{\min}$, $A \geq A_{\min}$, $DL \leq DL_{\max}$. Для достижения целевых значений планируются: довключение аудита на оставшихся серверах, контроль синхронизации времени, расширение обогащения активов, регулярный контроль WEF-подписок.

Заключение

В работе предложена и апробирована методика оценки качества логирования событий информационной безопасности на основе интегрального показателя дефицита логирования *DL* и итерационного цикла PDCA. В отличие от частных подходов к оценке *log completeness* и семантической корректности логов, показатель *DL* объединяет полноту поступления ожидаемых событий, валидность записей и межитерационную динамику устранения нарушений в единой количественной модели, ориентированной на применение в SOC/SIEM-контуре.

Методика основана на формировании профиля логирования, включающего матрицу требований **R**, набор обязательных полей *P**, период анализа *T* и пороговые значения $C_{\min}, A_{\min}, DL_{\max}$. Такой профиль позволяет оценивать не только факт поступления ожидаемых событий, но и пригодность полученных записей для последующей нормализации, корреляции и анализа.

На демонстрационном кейсе WEF/WEC показано снижение *DL* с 0,30 до 0,11 за пять итераций PDCA при росте числа корректных записей с 700 до 890. Полученные результаты подтверждают применимость методики для количественного контроля качества логирования, выявления причин дефицитов и оценки эффекта корректирующих мероприятий. В дальнейшем планируется развить модель за счёт вероятностного эталона потока событий, показателя пригодности событий к корреляции и сценарно-риск-взвешенной агрегации.

Литература

1. Сидак А.А. Стандартизация процессов регистрации событий безопасности, мониторинга и управления инцидентами информационной безопасности / А.А. Сидак, Д.А. Сидак // Двойные технологии. – 2023. – № 3 (104). – С. 58–62.

2. Канев А.Н. Мониторинг событий и обнаружение инцидентов информационной безопасности с использованием SIEM-систем // Международный студенческий научный вестник. – 2015. – № 3-1. – С. 122–123.

3. Cost of a Data Breach Report 2023 / IBM Security; Ponemon Institute. – 2023. – 85 p. [Электронный ресурс]. – URL: https://www.ibm.com/reports/data-breach (дата обращения: 10.03.2025).

4. Cisco Talos Incident Response. Quarterly Report: Q1 2024 Trends. – 2024 [Электронный ресурс]. – URL: https://blog.talosintelligence.com/talos-ir-q1-2024-quarterly-report (дата обращения: 12.03.2025).

5. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection – Information security management systems – Requirements. – Geneva: International Organization for Standardization, 2022. – 33 p.

6. Pascoe C. The NIST Cybersecurity Framework (CSF) 2.0 / C. Pascoe, S. Quinn, K. Scarfone. – Gaithersburg, MD: National Institute of Standards and Technology, 2024. – 32 p. (NIST CSWP 29). DOI: 10.6028/NIST.CSWP.29.

7. European Union Agency for Cybersecurity (ENISA). Technical Implementation Guidance on Cybersecurity Risk Management Measures. June 2025, version 1.0. – Luxembourg: Publications Office of the European Union, 2025. – 170 p. DOI:10.2824/2702548.

8. Kent K. Guide to Computer Security Log Management / K. Kent, M. Souppaya. – Gaithersburg, MD: National Institute of Standards and Technology (Special Publication 800-92), 2006. – 64 p.

9. Landauer M. Deep learning for anomaly detection in log data: A survey / M. Landauer, S. Onder, F. Skopik, M. Wurzenberger // Machine Learning with Applications. – 2023. – Vol. 12. – Art. 100470. DOI: 10.1016/j.mlwa.2023.100470.

10. Chuvakin A.A. Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management / A.A. Chuvakin, K.J. Schmidt, C. Phillips. – Burlington, MA: Syngress, 2013. – 460 p. DOI: 10.1016/C2010-0-65241-2.

11. ГОСТ Р 59548–2022. Защита информации. Регистрация событий безопасности. Требования к регистрируемой информации. – М.: Стандартинформ, 2022. – 70 с.

12. Жемчугов А.М. Цикл PDCA Деминга. Современное развитие / А.М. Жемчугов, М.К. Жемчугов // Проблемы экономики и менеджмента. – 2016. – № 2 (54). – С. 3–28.

13. ISACA. COBIT 2019 Framework: Governance and Management Objectives. – Schaumburg, IL: Information Systems Audit and Control Association, 2018. – 262 p.

14. Cichonski P. Computer Security Incident Handling Guide / P. Cichonski, T. Millar, T. Grance, K. Scarfone. – Gaithersburg, MD: National Institute of Standards and Technology, 2012. – 79 p. (NIST Special Publication 800-61 Rev. 2). DOI: 10.6028/NIST.SP.800-61r2.

15. Gujarati D.N. Basic Econometrics. – 4th ed. – New York: McGraw-Hill, 2003. – 1002 p.

16. Gardner E.S. Exponential smoothing: the state of the art // Journal of Forecasting. – 1985. – Vol. 4, No. 1. – P. 1–28. DOI: 10.1002/for.3980040103.

17. Nise N.S. Control Systems Engineering. – 8th ed. – Hoboken, NJ: Wiley, 2019. – 800 p.

18. Макаренко С.И. Аудит информационной безопасности: основные этапы, концептуальные основы, классификация мероприятий // Системы управления, связи и безопасности. – 2018. – № 1. – С. 1–29. DOI: 10.24411/2410-9916-2018-10101.

Иванов Андрей Валерьевич
Канд. техн. наук, доцент, зав. каф. защиты информации Новосибирского государственного технического университета (НГТУ)
К. Маркса пр-т, 20, г. Новосибирск, Россия, 630073
ORCID: 0000-0003-2002-8572
Тел.: +7 (383-3) 46-08-53, 8-913-708-10-00
Эл. почта: andrej.ivanov@corp.nstu.ru

Киселев Максим Алексеевич
Аспирант каф. защиты информации НГТУ
К. Маркса пр-т, 20, г. Новосибирск, Россия, 630073
ORCID: 0009-0001-2100-0611
Тел.: +7 (383-3) 46-08-53, 8-951-161-30-45
Эл. почта: asuszenfone_2@mail.ru

Поступила в редакцию: 23.01.2026.
Принята к публикации: 04.05.2026.

Ivanov A.V., Kiselev M.A.

Methodology for Assessing the Quality of Information Security Event Logging in an Iterative Management Cycle Using the Logging Deficit Metric

Relevance. The quality of information security event logging determines the completeness of subsequent normalization, correlation, and incident analysis in SIEM systems. Incomplete or structurally incorrect events reduce infrastructure observability and complicate incident response. ***Purpose of the study*** is to develop a methodology for the quantitative assessment of information security event logging quality based on the integral logging deficit indicator *DL* and the iterative PDCA cycle. ***Methods***. The study uses logging profile formalization, a matrix of event logging requirements, a set of mandatory fields, logging completeness and accuracy metrics, a recurrent *DL* dynamics model, and an IDEF0 representation of the methodology. ***Novelty***. A methodology is proposed that combines a logging requirements matrix, a set of mandatory fields, the integral *DL*-indicator, and inter-iteration accounting of newly detected and eliminated deficits in the PDCA cycle. ***Results***. Using a Windows authentication event stream as an example, a decrease in *DL* from 0.30 to 0.11 over five iterations is shown, with the number of valid records increasing from 700 to 890. ***Practical significance.*** The methodology can be applied for formalized logging quality control, prioritizing corrective actions, and integrating event quality assessment into regular information security monitoring processes.

Keywords: PDCA, SIEM, security event logging quality assessment, logging completeness and accuracy, logging deficit, mathematical model, iterative methodologies, ISO 27001, NIST CSF.
DOI: 10.21293/1818-0442-2026-29-1-114-123

References

1. Sidak A.A., Sidak D.A. Standardization of processes for security event registration, monitoring and information security incident management. Dual Technologies, 2023, no. 3(104), pp. 58–62 (in Russ.).

2. Kanev A.N. Monitoring sobytiy i obnaruzhenie intsidentov informatsionnoy bezopasnosti s ispol'zovaniem SIEM-sistem [Monitoring of events and detection of information security incidents using SIEM systems]. Mezhdunarodnyi studencheskii nauchnyi vestnik, 2015, no. 3-1, pp. 122–123 (in Russ.).

3. IBM Security; Ponemon Institute. Cost of a Data Breach Report 2023. IBM Security, 2023. 85 pp. Available online: https://www.ibm.com/reports/data-breach (accessed: 10 March 2025).

4. Cisco Talos Incident Response. Quarterly Report: Q1 2024 Trends. Cisco Talos, 2024. Available online: https://blog.talosintelligence.com/talos-ir-q1-2024-quarterly-report (accessed: 12 March 2025).

5. International Organization for Standardization. ISO / IEC 27001:2022. Information security, cybersecurity and privacy protection – Information security management systems. – Requirements. Geneva, ISO, 2022, 33 p.

6. Pascoe C., Quinn S., Scarfone K. The NIST Cybersecurity Framework (CSF) 2.0. Gaithersburg, MD, National Institute of Standards and Technology, 2024, 32 p. DOI: 10.6028/NIST.CSWP.29.

7. European Union Agency for Cybersecurity (ENISA). Technical Implementation Guidance on Cybersecurity Risk Management Measures. June 2025, version 1.0. Luxembourg, Publications Office of the European Union, 2025, 170 p. DOI: 10.2824/2702548.

8. Kent K., Souppaya M. Guide to Computer Security Log Management. Gaithersburg, MD, National Institute of Standards and Technology, 2006, 64 p. Special Publication 800-92.

9. Landauer M., Onder S., Skopik F., Wurzenberger M. Deep learning for anomaly detection in log data: A survey. Machine Learning with Applications, 2023, vol. 12, art. 100470. DOI: 10.1016/j.mlwa.2023.100470.

10. Chuvakin A.A., Schmidt K.J., Phillips C. Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management. Burlington, MA, Syngress, 2013, 460 p. DOI: 10.1016/C2010-0-65241-2.

11. GOST R 59548–2022. Zashchita informatsii. Registratsiya sobytiy bezopasnosti. Trebovaniya k registriruemoy informatsii [Information security. Security event logging. Requirements for logged information]. Moscow, Standartinform Publ., 2022, 70 p. (in Russ.).

12. Zhemchugov A.M., Zhemchugov M.K. Tsikl PDCA Deminga. Sovremennoe razvitie [Deming's PDCA cycle: modern development]. Problemy ekonomiki i menedzhmenta, 2016, no. 2 (54), pp. 3–28 (in Russ.).

13. ISACA. COBIT 2019 Framework: Governance and Management Objectives. Schaumburg, IL, Information Systems Audit and Control Association, 2018, 262 p.

14. Cichonski P., Millar T., Grance T., Scarfone K. Computer Security Incident Handling Guide. Gaithersburg, MD, National Institute of Standards and Technology, 2012, 79 p. NIST Special Publication 800-61 Rev. 2. DOI: 10.6028/NIST.SP.800-61r2.

15. Gujarati D.N. Basic Econometrics. 4th ed. New York, McGraw-Hill, 2003, 1002 p.

16. Gardner E.S. Exponential smoothing: the state of the art. Journal of Forecasting, 1985, vol. 4, no. 1, pp. 1–28. DOI: 10.1002/for.3980040103.

17. Nise N.S. Control Systems Engineering. 8th ed. Hoboken, NJ, Wiley, 2019, 800 p.

18. Makarenko S.I. Audit of Information Security – the Main Stages, Conceptual Framework, Classification of Types. Systems of Control, Communication and Security, 2018, no. 1, pp. 1–29. DOI: 10.24411/2410-9916-2018-10101 (in Russ.).

Andrey V. Ivanov
Candidate of Technical Sciences, Associate Professor,
Head of the Department of Information Security,
Novosibirsk State Technical University (NSTU)
20, K. Marx Ave., Novosibirsk, Russia, 630073
ORCID: 0000-0003-2002-8572
Phone: +7 (383-3) 46-08-53, 8-913-708-1000
Email: andrej.ivanov@corp.nstu.ru

Maxim A. Kiselev
Postgraduate Student, Department of Information Security, NSTU
20, K. Marx Ave., Novosibirsk, Russia, 630073
ORCID: 0009-0001-2100-0611
Phone: +7 (383-3) 46-08-53, 8-951-161-3045
Email: asuszenfone_2@mail.ru

Received: 23.01.2026.
Accepted: 04.05.2026.