ОБЪЯСНИМАЯ ИНТЕРПРЕТАЦИЯ ИНЦИДЕНТОВ НА ОСНОВЕ БОЛЬШОЙ ЯЗЫКОВОЙ МОДЕЛИ И МЕТОДА ГЕНЕРАЦИИ С ДОПОЛНЕННОЙ ВЫБОРКОЙ

Котенко И. В.[1], Абраменко Г. Т.[2]

DOI: 10.21681/2311-3456-2025-5-58-67

Цель исследования: *повысить достоверность и объяснимость интерпретации оповещений системы обнаружения и предотвращения вторжений Suricata за счет онтологически обогащенного графа знаний, гетерогенных графовых представлений и метода генерации с дополненной выборкой (Retrieval-Augmented Generation, RAG) на основе локальной большой языковой модели (Large Language Model, LLM).*

Методы исследования: *построение онтологически управляемого графа знаний, связывающего данные Suricata с тактиками/техниками MITRE ATT&CK; обучение гетерогенной графовой нейросети (HGNN) для получения контекстных векторных представлений узлов; извлечение релевантного контекста по ближайшим соседям в пространстве эмбеддингов; генерация объяснений через локальную LLM (с 7B параметрами) по RAG-конвейеру; экспериментальная оценка на корпусе из 25000 оповещений Suricata с использованием метрик точности интерпретации, доли галлюцинаций и релевантности.*

Полученные результаты: *разработан онтологически управляемый метод интерпретации оповещений Suricata, обеспечивающий более полное и корректное объяснение по сравнению с базовым подходом. Показано, что использование онтологий позволяет повысить содержательность объяснений на 15 % при несущественном увеличении времени генерации. В результате, интеграция онтологии и гетерогенного графа знаний существенно повышает корректность сопоставления оповещений об инцидентах безопасности с техниками MITRE ATT&CK и снижает риск неверных объяснений.*

Научная новизна: *предложена интеграция онтологии и гетерогенного графа знаний с RAG-генерацией поверх локальной LLM для привязки низкоуровневых событий системы обнаружения и предотвращения вторжений к техникам MITRE ATT&CK; показана применимость онтологий к задачам объяснимого ИИ для кибербезопасности.*

Ключевые слова: *кибербезопасность; системы обнаружения и предотвращения вторжений; поиск угроз безопасности; объяснимый искусственный интеллект; граф знаний; глубокое обучение: гетерогенная графовая нейросеть; MITRE ATT&CK; LLM; RAG; Suricata.*

1. Введение

Современные центры мониторинга кибербезопасности (Security Operations Center, SOC) ежедневно сталкиваются с лавинообразным потоком сетевых событий и оповещений систем обнаружения и предотвращения вторжений (Intrusion Detection and Prevention System, IDPS), таких как Suricata[3]. Значительная часть этих оповещений оказывается нерелевантной или малоинформативной для оперативного реагирования, а попытка вручную соотнести низкоуровневые сигнатуры с высокоуровневыми сценариями атак приводит к перегрузке аналитиков и задержкам при принятии решений [1]. Таким образом, повышается актуальность методов, которые не только выявляют подозрительную активность, но и интерпретируют смысл каждого срабатывания в терминах тактик, техник и процедур злоумышленника, понятных специалисту SOC.

Попытка напрямую использовать большие языковые модели (Large Language Model, LLM) для построения таких объяснений выглядит естественной: они умеют излагать сложные технические детали на естественном языке. Однако LLM склонны к фактологическим ошибкам и «галлюцинациям», что в критичных сценариях кибербезопасности недопустимо [2, 3].

Для решения указанных рисков используется генерация с дополненной выборкой (Retrieval-Augmented Generation, RAG), когда языковая модель получает заранее отобранные релевантные сведения из внешнего источника знаний и генерирует формулировки, используя эти сведения [4]. В гибридных архитектурах к RAG добавляются структурированные представления и графовые нейронные сети (Graph Neural Network, GNN), что повышает точность и воспроизводимость рассуждений [5].

1 Котенко Игорь Витальевич, доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник и руководитель лаборатории проблем компьютерной безопасности, ФГБУН «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН), г. Санкт-Петербург, Россия. E mail: ivkote@comsec.spb.ru

2 Абраменко Георгий Тимофеевич, аспирант факультета безопасности информационных технологий, ФГАОУ ВО «Национальный исследовательский университет ИТМО», г. Санкт-Петербург, Россия. E-mail: gtabramenko@itmo.ru

3 Suricata – Intrusion Detection and Prevention System. https://suricata.io

Наиболее перспективным источником такого знания для задач SOC являются формальные онтологии и графы знаний (Knowledge Graph, KG) по кибербезопасности (например, MITRE ATT&CK, CAPEC и др.)[4]. Использование подхода KG-Infused RAG [6] позволит связать оповещения IDPS с тактиками и техниками атак, индикаторами компрометации, уязвимостями и инструментами [7, 8] на основе стандартизированного словарь и практик проактивного поиска угроз [9]. В [10] демонстрируются промышленные решения корреляции инцидентов с учетом миллиардов событий, опирающиеся на графовые структуры.

Вместе с тем остается недостаточно изученным вопрос, связанный с тем, насколько гетерогенная структура графа, включающая различные типы узлов и связей (например, узлы типа Оповещение, Сигнатура, Техника, Индикатор Alert, Signature, Technique, Indicator и т.д.) и явное онтологическое ядро знаний (ATT&CK/CAPEC) улучшают качество именно интерпретаций, а не только обнаружения. Это особенно важно для оповещений IDPS в реальных потоках данных [11, 12], тогда как формализованное моделирование угроз на базе MITRE ATT&CK[5] влияет на пригодность объяснений для аудита и действий по реагированию на инциденты [13].

В данной работе развивается направление, связанное с обозначенными перспективными направлениями в исследованиях, и предлагается онтологически управляемый метод интерпретации оповещений Suricata на основе гетерогенного графа знаний и LLM с использованием RAG. Ключевая идея – представить каждое оповещение как узел графа, связанный с узлами сигнатур, техник и тактик из онтологии атак; далее обученная гетерогенная графовая нейронная сеть (Heterogeneous GNN, HGNN) вычисляет векторное представление события и извлекает ближайший контекст из графа, который затем используется LLM для достоверного объяснения. Такой подход обеспечивает отслеживаемость утверждений до конкретных сущностей базы знаний, снижение доли вымышленных фактов, повышение полезности текста для принятия решений аналитиком SOC.

2. Анализ релевантных работ

Проблема устранения избыточных и ложных оповещений, а также приоритизации инцидентов в SOC является актуальным направлением исследований в области кибербезопасности. Так, в [1] приводятся данные исследования SOC, в котором отмечается, что в проанализированных предприятиях обрабатывается свыше 2,4 млн оповещений ежедневно, причем до 70 % из них являются ложными. Предлагаются подходы к интеллектуальной приоритизации оповещений с помощью больших языковых моделей, интегрированных в системы фильтрации, оценки и категоризации событий.

Некоторые публикации фокусируются на объяснимости интеллектуальных моделей для классификации сетевых атак. В частности, в [2] оценивались методы объяснимого искусственного интеллекта (ИИ) (Explainable AI, XAI), и было указано, что недостаток прозрачности мешает практическому применению моделей машинного обучения (Machine Learning, ML) для обнаружения угроз.

В [14] представлена диалоговая система, функционирующая на основе агрегированных баз знаний по уязвимостям и инцидентам, в которой сначала извлекаются подтвержденные факты, а затем формируется текстовое пояснение. Такой подход демонстрирует рост точности ответов при типичных запросах аналитика. В [15] источником контекста служат одновременно два графа: внутренний граф событий и журналов организации и внешний граф знаний по кибербезопасности. Из каждого графа выбираются согласованные фрагменты, и уже на их основе формируются пояснения.

В [16] для задач реагирования на инциденты предлагается подход, где структурированные запросы к внешним источникам сведений об угрозах сочетаются с семантическим поиском по локальным материалам. В результате реализации такого подхода формируется контекст, который используется для выдачи конкретных рекомендаций по реагированию на инциденты.

Отдельно развивается направление объединения графовых моделей с генеративным ИИ. В [17] представлена технология корреляции событий безопасности, основанная на графах знаний, в которых события связываются через уязвимости, активы и тактики атак, что обеспечивает выявление сложных атак. Для анализа данных из журналов безопасности предложены методы автоматизированного построения графов [18] и виртуальные графы знаний для анализа логов из разнородных источников [19]. Такой подход снижает затраты на предварительную интеграцию данных и повышает пригодность данных для последующего анализа. Жизнеспособность графовых структур подтверждается в системах корреляции инцидентов, работающих с большими объемами событий [10].

Заметный прогресс в исследованиях по системам обнаружения вторжений связан с переходом от классических признаков к графовым представлениям. В [20, 21] показано, что явный учет связей между объектами инфраструктуры и событиями повышает

4 Syed Z., Padia A., Finin T., Mathews L., Joshi A. UCO: A Unified Cybersecurity Ontology // Proc. AAAI Workshop on Artificial Intelligence for Cyber Security. 2016. 8 p.

5 Strom, B.; et al. MITRE ATT&CK: Design and Philosophy. MITRE Technical Report, 2018.

полноту и устойчивость обнаружения сложных атак. Также развиваются методы реализации объяснимости графовых моделей. Такие методы позволяют определять, какие вершины графа и какие ребра повлияли на итоговое решение. Кроме того, вводятся количественные показатели устойчивости объяснений [22]: проверяется, сохраняется ли интерпретация при небольших изменениях входных данных и структуры графа, а также при изменении статистики данных между обучением и использованием.

При реализации подходов, основанных на RAG, складывается общее понимание, что качество модуля поиска, полнота охвата и согласованность найденных фактов определяют применимость и полезность итогового ответа. В [5, 23, 24] предлагаются решения, в которых сам граф знаний направляет поиск и задает допустимые цепочки рассуждений, что дает возможность повысить согласованность и уменьшить риск вымышленных утверждений.

Подобные системы демонстрируют, что интеграция внешних знаний способна уменьшить галлюцинации моделей и повысить точность объяснений. Для объяснения оповещений IDPS используются подходы, основанные на комбинировании гетерогенных графов и LLM. Так в [25] представлено использование XG-NID с HGNN и больших языковых моделей для объяснения сетевых вторжений, а также предлагается подход к интерпретации решений GNN в кибербезопасности.

В отечественных исследованиях последних лет фиксируется устойчивый тренд применения методов искусственного интеллекта к задачам кибербезопасности. Развиваются подходы к федеративному обучению для систем обнаружения вторжений, например, в [26] представлена архитектура федеративной IDS/IDPS с экспериментальной валидацией на репрезентативных данных. Отдельное направление составляет атрибуция целевых атак и профилирование нарушителей, например, в [27] систематизированы модели, источники данных и методики, определены ограничения и требования к системам атрибуции атак. Параллельно формируется графовый аппарат оценки защищенности предложена модель оценки на графе эксплойтов, обеспечивающая количественные показатели для принятия решений [28]. Развитие идей федеративного обучения демонстрирует предложенная в [29] распределенная система обнаружения сетевых атак, сочетающая федеративное и трансферное обучение. Для систем обнаружения вторжений в кибер-физических системах обобщены источники данных, методы и метрики, выявлены ограничения наборов данных и типовые ошибки постановки экспериментов [30]. Для сетей Интернета вещей применяются многозадачное обучение и гибридные методы формирования выборок, повышающие полноту при выраженном дисбалансе классов [31]. Для автоматизации и приоритизации реагирования на инциденты разработаны методы прогнозирования категорий уязвимостей по конфигурациям устройств, что позволяет подготовить признаки для ранжирования задач реагирования на инциденты [32].

Обобщая рассмотренные исследования, следует отметить, что эффективная интерпретация оповещений IDPS требует использования формализованных знаний (онтологии и графы знаний). Применение гетерогенных графовых моделей для представления разнотипных сущностей и связей, а также использование RAG-подходов для обеспечения текстовых объяснений проверяемых фактов, извлеченных из внешних баз знаний, потенциально сможет уменьшить вероятность недостоверных утверждений и повысить практическую ценность выводов для аналитиков SOC. Вместе с тем недостаточно проработанными остаются вопросы построения интегрированного конвейера, в рамках которого оповещения для IDPS систем системно встраиваются в онтологически обогащенный гетерогенный граф и далее используются в качестве внешнего хранилища знаний для языковой модели при формировании кратких и проверяемых пояснений по каждому оповещению.

3. Варианты архитектур системы объяснения оповещений

В статье представлено три варианта предлагаемой архитектуры системы объяснения оповещений IDPS, основанных на комбинировании поиска информации на графе знаний с генеративными возможностями LLM.

Каждая из архитектур представлена на (рис. 1).

Сопоставление трех архитектур, представленное в (табл. 1), показывает, что усложнение модели памяти от однородного графа к гетерогенному и далее к онтологически управляемому графу последовательно снижает риск шумовых (ложных) соответствий и повышает «проверяемость» объяснений. Это выражается в переходе от выбора единственного «наиболее важного» пути (GNN-RAG) к формированию согласованного подграфа доказательств (HGNN-RAG) и, наконец, к отбору контекста под жесткими онтологическими ограничениями с трассировкой ссылок на узлы ATT&CK/CAPEC (Ontology-KG-HGNN-RAG).

При этом возрастает стоимость подготовки данных: если базовому варианту достаточно минимальной связности KG, то онтологически управляемый вариант требует поддержания соответствий сигнатур и событий элементам онтологии и регулярного обновления отображений (маппингов).

С учетом табл. 1 можно сформулировать следующие выводы.

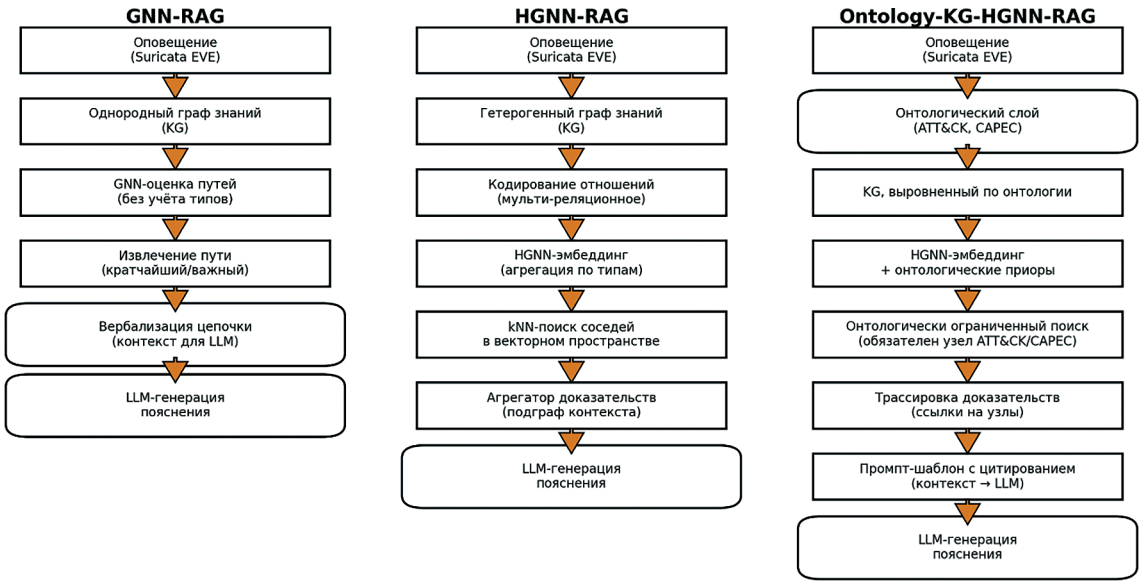

Рис. 1. Архитектуры GNN-RAG, HGNN-RAG, Ontology-KG-HGNN-RAG

Во-первых, GNN-RAG целесообразен как стартовый или резервный режим: он прост в развертывании и дает минимально необходимую интерпретируемость, но чувствителен к косвенным связям и потому хуже подходит для отчетных сценариев SOC.

Во-вторых, HGNN-RAG является рабочим компромиссом: он учитывает природу сущностей и взаимосвязей, формирует более чистый контекст и обеспечивает устойчивый прирост качества без радикального роста операционных затрат.

В-третьих, Ontology-KG-HGNN-RAG приоритетен для сред с повышенными требованиями для аудита: он обеспечивает комплекс ссылок на онтологию и минимизирует недостоверные утверждения, однако требует строгости во ведении онтологического слоя и подтверждения маппингов.

С точки зрения эксплуатационных характеристик, рост точности и релевантности при переходе к онтологически управляемой архитектуре сопровождается увеличением стоимости предлагаемого решения. Необходимо поддерживать онтологическую базу (обновления ATT&CK/CAPEC, ревизии соответствий сигнатур) в актуальном состоянии, а также периодически переобучать HGNN для учета новых типов

Таблица 1.

Сопоставление архитектур по ключевым характеристикам

Характеристика	GNN-RAG	HGNN-RAG	Ontology-KG-HGNN-RAG
Тип графа	Однородный KG	Гетерогенный KG (типы узлов/связей)	Гетерогенный KG + явное онтологическое ядро
Выбор контекста	Кратчайший/значимый путь; kNN	Мульти-реляционный агрегатор; kNN	Онтологически ограниченный поиск; kNN
Прослеживаемость к ATT&CK/CAPEC	Косвенная	Частичная	Гарантированная (узлы-опоры в каждой цепочке)
Риск шумовых связей	Повышенный	Сниженный	Минимизирован за счет онтологических ограничений
Ожидаемая доля «галлюцинаций» LLM	Выше	Ниже	Наименьшая
Практический вывод	Базовая интерпретация	Семантически богаче	Наиболее строгие и точные объяснения

связей. В обмен на эти затраты формируются объяснения, пригодные для требуемой отчетности и повторной проверки: каждое утверждение привязывается к явно идентифицируемым узлам знания, а отбор контекста является воспроизводимым.

4. Эксперименты

4.1. Описание экспериментов

Эксперименты выполнялись на воспроизводимом стенде SOC. В качестве сетевой основы использовался корпус CIC-IDS2017[6] после его обработки на Suricata (Emerging Threats, релиз 2024.1). Получено около 25000 срабатываний примерно по пяти десяткам сигнатур. Чтобы сравнение трех архитектур (и соответствующих конвейеров) было однозначным, было отобрано 30 репрезентативных правил, и для одних и тех же событий формировались объяснения на основе использования трех выбранных подходов (GNN-RAG, HGNN-RAG, Ontology-KG-HGNN-RAG) – всего 90 текстов.

Ниже приведен фрагмент EVE-события, на основе которого строится граф и формируется контекст для LLM:

```
{
   "timestamp": "2025-05-26T14:07:31.342Z",
   "src_ip": "10.24.5.12", "src_port": 53342,
   "dest_ip": "8.8.8.8",  "dest_port": 53,
   "proto": "UDP",
   "alert": {
      "signature_id": 1234567,
      "signature": "ET DNS Possible DNS
Tunneling",
      "category": "A Network Trojan was
detected",
      "severity": 2
   }
}
```

Сопоставление полей EVE с онтологией/графом выполнено так: Signature → Technique (ATT&CK T1071), Technique → Tactic (Command-and-Control), Event → Indicator (DNS/53-UDP).

4.2. Архитектура конвейера и параметры эксперимента

Все три конвейера имеют одинаковый режим обучения графовой части: двухслойная модель с размерностью представлений $d = 64$ и контрастивным обучением на парах «позитив/негатив» (позитивы – вручную размеченные связи «Событие → Техника» и «Событие → ПО/инструмент», негативы – случайные несвязанные пары). В процессе вывода для узла «событие» выбираются $k = 3$ ближайших факта из графа по косинусному сходству (формула (1)) и из них собирается текстовый контекст для LLM.

$$sim(u,v) = \frac{(h_v^T h_u)}{\|h_v\|_2 \cdot \|h_u\|_2};\ sim \in [-1;1], \qquad (1)$$

где: u – узел «событие», v – узел «факт»; h_v, $h_u \in R^d$ – их векторные представления; $\| \ \|_2$ – евклидова норма.

Чем значение sim ближе к 1, тем больше доверие, что событие и найденный факт описывают один и тот же контекст; по убыванию sim берутся k лучших соответствий.

Во второй архитектуре (HGNN-RAG) эмбеддинги считаются с учетом типов ребер: вклад переходов «Signature → Technique» и «Technique → Tactic» различается. Обновление представлений задается формулой (2), где сумма берется по всем типам $r \in R$ и по соседям соответствующего типа:

$$h_i^{(l+1)} = ReLU\left(W_{self}\, h_i^{(l)} + \sum_{r \in R} \sum_{j \in N_r(i)} \frac{1}{c_{i,r}} W_r h_j^{(l)} \right), \qquad (2)$$

где: i – узел, l – номер слоя; R – типы связей (например, «Сигнатура → Техника», «Техника → Тактика», «Событие → Индикатор»); N_r – соседи i по связи типа r; W_{self}, $W_r \in \mathbb{R}^{d \times d}$ – обучаемые матрицы; $c_{i,r} = \max(1, |N_r(i)|)$ – нормировка вклада «плотных» узлов.

Учет типов связей через W_r уменьшает число неверных соответствий по сравнению с однородным графом.

В онтологически управляемом конвейере (Ontology-KG-HGNN-RAG) вводится правило: каждая цепочка обоснования должна проходить через узел ATT&CK или CAPEC. Это обеспечивает проверяемость утверждений и устраняет выводы вне фактического контекста события.

Генерация выполнялась на локальной LLM (с примерно 7B параметрами); среднее время ответа 3–5 с (поиск контекста + генерация).

4.3. Метрики и процедура оценки

Оценка проводилась на одном и том же множестве событий, что обеспечивает сопоставимость трех модификаций конвейера. Эталонная разметка основной техники выполнялась независимыми экспертами центра мониторинга безопасности; для исключения переобучения использовалась отложенная выборка. Порог включения факта в контекст – фиксирован и одинаков для всех вариантов.

Ключевая метрика (основная техника):

$$\text{Precision@1} = \frac{1}{N} \sum_{i=1}^{N} 1\{\hat{t}_i = t_i^*\}, \qquad (3)$$

где N – число проверяемых событий; $\hat{t}_i$ – идентификатор основной техники ATT&CK, указанной в объяснении; t_i^* – эталон; $1\{\}$ – индикаторная функция.

Доверительные интервалы для средних экспертных оценок считались по стандартной формуле t-интервала; согласованность мнений проверялась коэффициентом координации Кендэлла W.

6 https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset

Таблица 2.

Качество генерируемых объяснений на 30 сигнатурах (90 текстов)

Метрика / Подход	GNN-RAG	HGNN-RAG	Ontology-KG-HGNN-RAG
Precision@1 (техника)	0,60	0,73	0,83
Recall (ключевые факты)	0,65	0,80	0,92
Длина объяснения, слов	33 ± 7	41 ± 6	46 ± 5
BLEU-2 (10 эталонов)	0,47	0,53	0,58
Корректность (эксперты)	3,4 ± 0,3	4,2 ± 0,2	4,8 ± 0,1
Полнота (эксперты)	3,1 ± 0,4	3,9 ± 0,3	4,6 ± 0,2
Ясность (эксперты)	3,7 ± 0,3	4,4 ± 0,2	4,8 ± 0,1
Полезность (эксперты)	3,0 ± 0,4	4,1 ± 0,3	4,7 ± 0,2

4.4. Результаты экспериментов

Сопоставление трех модификаций конвейера на одном и том же множестве событий показало устойчивое преимущество гетерогенного графа с онтологическим ограничением. По ключевой метрике Precision@1 этот вариант уверенно превосходит однородный граф и опережает «просто гетерогенный»; прирост качества сопровождается умеренным увеличением времени ответа, что приемлемо для практической эксплуатации центра мониторинга безопасности.

Сводные результаты представлены в табл. 2.

Показатель BLEU-2 растет в пользу вариантов с учетом типов связей, что отражает лучшее соответствие эталонным описаниям. Экспертные оценки пяти аналитиков SOC демонстрируют согласованное преимущество по «Корректности», «Полноте», «Ясности» и «Полезности».

GNN-RAG обеспечивает наименьшее время обработки, но склонен к «общим» пояснениям без четкой привязки к технике, а HGNN-RAG снижает долю неверных привязок, но при скудном контексте все еще может «останавливаться» на уровне тактики. Ontology-KG-HGNN-RAG повышает проверяемость выводов: каждое ключевое утверждение проходит через узел онтологии, а трассировка по идентификаторам в тексте объяснения позволяет быстро сверить вывод с первоисточником.

4.5. Анализ результатов и типичные ошибки

Базовый GNN-RAG дал корректную, но расплывчатую формулировку без привязки к ATT&CK. HGNN-RAG точно указал на T1568.002 (DGA). Ontology-KG-HGNN-RAG дополнительно связал событие с тактикой Command-and-Control и предложил проверку хоста на вредоносные DGA – эксперты признали этот вариант наиболее полезным.

Чтобы проиллюстрировать различия в поведении подходов, приведем два характерных примера.

ET SCAN SMB1. В базовом варианте встречались выводы «сверх контекста» (например, упоминание потенциального удаления файлов без опоры в графе). В онтологической версии такие домыслы исчезли, поскольку отбор контекста обязан проходить через узлы ATT&CK/CAPEC.

Сигнатура с недостаточным контекстом «ET POLICY PE EXE or DLL Windows file download». Срабатывание фиксирует загрузку PE по HTTP/HTTPS и само по себе не доказывает атаку. Однородный и гетерогенный графы при дефиците данных обычно выдают тактику без конкретной техники. Онтологически управляемый вариант принудительно проводит контекст через MITRE ATT&CK и дает T1105 («Передача инструментария») с проверяемыми признаками. Итоговое объяснение содержит технику и обоснование, пригодное для эскалации.

Основные источники ошибок. Основные источники ошибок во всех вариантах связаны с дефицитом контекста и неполнотой соответствий «сигнатура → техника». На практике это устранимо регулярной актуализацией справочника соответствий, обогащением графа проверенными связями и явным правилом «нет достаточных оснований – техника не указывается», чтобы избегать повторной интерпретации.

4.6. Ограничения и практические замечания

Качество результатов зависит от полноты и актуальности онтологического слоя. При появлении новых техник и описаний атак требуется оперативное обновление соответствий и переиндексация графа. В противном случае верхняя планка качества ограничивается отсутствием опорных узлов. Источники данных также накладывают ограничения: часть событий получена на основе открытого набора трафика и может отличаться от специфики конкретной организации. Для переноса в промышленную среду полезно дополнительно проверять систему на реальных журналах предприятия.

По ресурсам система остается пригодной для повседневной работы SOC: рост времени ответа при переходе к онтологическому варианту является умеренным, однако при масштабировании на значительно большие графы потребуется инженерная оптимизация — приближенные структуры для поиска ближайших соседей, шардирование графа и предфильтрация по приоритетам событий.

С точки зрения эксплуатации важно обеспечить трассируемость: в каждое объяснение включаются идентификаторы ATT&CK/CAPEC/SID, а интерфейс должен позволять переход к источнику соответствия. Рекомендуется регламентировать пороги включения фактов, периодичность обновления онтологии и порядок валидации изменений, а также хранить версионированные слои знаний для воспроизводимости результатов. В перспективе предполагается расширить проверку на других наборах открытого трафика и на реальных потоках из производственной сети, а также изучить влияние адаптивных порогов по классам событий.

Заключение

В работе предложен подход к объяснению оповещений IDPS (на примере Suricata) с использованием графовых знаний и больших языковых моделей.

Проведенное исследование позволяет сделать следующие выводы:

- Онтологии кибербезопасности (в частности, MITRE ATT&CK, CAPEC, STIX) содержат ценную информацию для интерпретации оповещений, генерируемых при мониторинге инцидентов безопасности. Их интеграция в процесс объяснения позволяет увязать разрозненные низкоуровневые события с общей схемой реализации атаки, что повышает осведомленность аналитика и качество решений.
- Разработан подход Ontology-KG-HGNN-RAG, сочетающий гетерогенную графовую нейронную сеть и онтологически обогащенный граф знаний для извлечения релевантного контекста, а также генеративную модель для формирования текстового объяснения. Данный подход превосходит базовый вариант (без учета типов узлов и онтологий) по полноте и корректности генерируемых объяснений примерно на 15–20 % по ключевым метрикам, а также получил более высокие оценки экспертов-практиков (на 1–1,5 балла по 5-балльной шкале).
- Показана реализуемость автоматической системы объяснений в условиях, приближенных к реальным: использованы открытые данные трафика и актуальные базы знаний. Таким образом, технология может найти применение на практике в составе систем оркестрации, автоматизации и реагирования на инциденты информационной безопасности (SOAR, Security Orchestration, Automation and Response) или систем управления событиями и информацией безопасности нового поколения (NG-SIEM, Next-Generation Security Information and Event Management).
- Ограничения метода включают зависимость от полноты базы знаний и необходимость поддержки в актуальном состоянии, а также ресурсоемкость при масштабировании на очень большие потоки. Требуется дальнейшая проработка механизмов фильтрации и реализации распределенных вычислений. Кроме того, остаются нерешенными вопросы интеграции оператора в цикл принятия решения – как максимально удобно представить объяснение и учесть обратную связь.
- Результаты свидетельствуют о перспективности объединения графовых знаний и LLM для задач кибербезопасности. Данный подход может быть расширен на смежные проблемы, включая объяснение решений систем обнаружения аномалий, обогащение уведомлений SIEM, автоматизированное написание отчетов об инцидентах на основе сырых логов и др.

В будущем планируется развивать данное направление, в том числе исследовать возможность адаптации модели к русскоязычным онтологиям, а также реализовать прототип плагина к популярным SIEM-системам, который на лету будет выводить «объяснительные подсказки» для каждого предупреждения. Предложенный подход может послужить основой для нового поколения интеллектуальных помощников аналитика SOC, совмещающих возможности ИИ и прозрачность знаний.

Благодарность. Работа выполнена при частичной финансовой поддержке бюджетной темы FFZF-2025-0016.

Рецензент: Липатников Валерий Алексеевич, доктор технических наук, профессор, заслуженный деятель науки Российской Федерации, старший научный сотрудник научно-исследовательского центра Военной академии связи имени Маршала Советского Союза С. М. Буденного, Санкт-Петербург, Россия. E-mail: lipatnikovanl@mail.ru

Литература

1. Singh A. Contextual Threat Intelligence and Alert Prioritization with Foundation-Sec-8B // International Journal of Artificial Intelligence Research and Development. 2025. Vol. 3, No. 1. P. 131–145. DOI: 10.34218/IJAIRD_03_01_009.
2. Arreche O., Guntur T., Abdallah M. XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems // Applied Sciences. 2024. Vol. 14, No. 10. Article 4170. DOI: 10.3390/app14104170.
3. Hassanin M., Moustafa N. A Comprehensive Overview of Large Language Models (LLMs) for Cyber Defences: Opportunities and Directions // arXiv preprint arXiv:2405.14487. 2024. DOI: 10.48550/arXiv.2405.14487.
4. Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q. A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models // Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'24). 2024. P. 6491–6501.
5. Mavromatis C., Karypis G. GNN-RAG: Graph Neural Retrieval for Large Language Model Reasoning // arXiv preprint arXiv:2405.20139. 2024. DOI: 10.48550/arXiv.2405.20139.
6. Wu D., Yan Y., Liu Z., Liu Z., Sun M. KG-Infused RAG: Augmenting Corpus-Based RAG with External Knowledge Graphs // arXiv preprint arXiv:2506.09542. 2025. DOI: 10.48550/arXiv.2506.09542.
7. Al-Sada B., Sadighian A., Oligeri G. MITRE ATT&CK: State of the Art and Way Forward // ACM Computing Surveys. 2024. Vol. 57, No. 1. Article 12. P. 1–37. DOI: 10.1145/3687300.
8. Li H., Shi Z., Pan C., Zhao D., Sun N. Cybersecurity Knowledge Graphs Construction and Quality Assessment // Complex & Intelligent Systems. 2024. Vol. 10. P. 1201–1217. DOI: 10.1007/s40747-023-01205-1.
9. Hemberg E., Kelly J., Shlapentokh-Rothman M., Reinstadler B., Xu K., Rutar N., O'Reilly U.-M. Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Platforms for Cyber Hunting // arXiv preprint arXiv:2010.00533. 2021. DOI: 10.48550/arXiv.2010.00533.
10. Freitas S., Gharib A. GraphWeaver: Billion-Scale Cybersecurity Incident Correlation // Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM'24). 2024. P. 4479–4486. DOI: 10.1145/3627673.3680057.
11. Sikos L. F. Cybersecurity Knowledge Graphs // Knowledge and Information Systems. 2023. Vol. 65, No. 9. P. 3511–3531. DOI: 10.1007/s10115-023-01860-3.
12. Zhao X., Jiang R., Han Y., Li A., Peng Z. A Survey on Cybersecurity Knowledge Graph Construction // Computers & Security. 2024. Vol. 136. Article 103524. DOI: 10.1016/j.cose.2023.103524.
13. Xiong W., Legrand E., Åberg O., Lagerström R. Cyber Security Threat Modeling Based on the MITRE Enterprise ATT&CK Matrix // Software and Systems Modeling. 2022. Vol. 21. P. 157–177. DOI: 10.1007/s10270-021-00898-7.
14. Arikkat D. R., Abhinav M., Binu N., Rajendran N., Bhattacharjee D., Das R. K., Ajay C. R., Shiyam Sundar R., Bhavesh S. IntelBot: A Retrieval-Augmented Large Language Model Chatbot for Cyber Threat Knowledge Delivery // arXiv preprint arXiv:2411.05442. 2024. DOI: 10.48550/arXiv.2411.05442.
15. Kurniawan K., Kiesling E., Ekelhart A. CyKG-RAG: Towards Knowledge-Graph Enhanced Retrieval-Augmented Generation for Cybersecurity // CEUR Workshop Proceedings. 2024. Vol. 3950. P. 51–64.
16. Tellache A., Amara-Korba A., Mokhtari A., Moldovan H., Ghamri-Doudane Y. Advancing Autonomous Incident Response: Leveraging LLMs and Cyber Threat Intelligence // arXiv preprint arXiv:2508.10677. 2025. DOI: 10.48550/arXiv.2508.10677.
17. Qi Y., Gu Z., Li A., Zhang X., Shafiq M., Mei Y., Lin K. Cybersecurity Knowledge Graph Enabled Attack Chain Detection for Cyber-Physical Systems // Computers & Electrical Engineering. 2023. Vol. 108. Art. 108660. DOI: 10.1016/j.compeleceng.2023.108660.
18. Ekelhart A., Ekaputra F.J., Kiesling E. SLOGERT: Automated Log Knowledge Graph Construction // The Semantic Web – ESWC 2021. Lecture Notes in Computer Science. 2021. Vol. 12731. P. 219–234. DOI: 10.1007/978-3-030-77385-4_16.
19. Kurniawan K., Ekelhart A., Kiesling E., Winkler D., Quirchmayr G., Tjoa A.M. Virtual Knowledge Graphs for Federated Log Analysis // Proceedings of the 16th International Conference on Availability, Reliability and Security (ARES'21). 2021. Article 50. 10 p. DOI: 10.1145/3465481.3470077.
20. Bilot T., El-Madhoun N., Al-Agha K., Zouaoui A. Graph Neural Networks for Intrusion Detection: A Survey // IEEE Access. 2023. Vol. 11. P. 49114–49139. DOI: 10.1109/ACCESS.2023.3275789.
21. Zhong M., Lin M., Zhang C., Xu Z. A Survey on Graph Neural Networks for Intrusion Detection Systems: Methods, Trends and Challenges // Computers & Security. 2024. Vol. 141. Article 103821. DOI: 10.1016/j.cose.2024.103821.
22. Fang J., Liu W., Gao Y., Liu Z., Zhang A., Wang X., He X. Evaluating Post-hoc Explanations for Graph Neural Networks via Robustness Analysis // Advances in Neural Information Processing Systems (NeurIPS'23), Oral. 2023.
23. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W.-t., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474. DOI: 10.5555/3495724.3495881.
24. Zhu X., Xie Y., Liu Y., Li Y., Hu W. Knowledge Graph-Guided Retrieval-Augmented Generation (KG²RAG) // Proceedings of NAACL 2025 (Long Papers). 2025. P. 8912–8924. DOI: 10.18653/v1/2025.naacl-long.449.
25. Farrukh Y. A., Wali S., Khan I., Bastian N. D. Xg-nid: Dual-modality network intrusion detection using a heterogeneous graph neural network and large language model // Expert Systems with Applications. 2025. p. 128089.
26. Новикова Е. С., Бухтияров М. А., Котенко И. В., Саенко И. Б., Федорченко Е. В. Обнаружение вторжений на основе федеративного обучения: архитектура системы и эксперименты // Вопросы кибербезопасности. 2023. № 6(58). С. 50–66. DOI: 10.21681/2311-3456-2023-6-50-66.
27. Котенко И. В., Хмыров С. С. Анализ моделей и методик, используемых для атрибуции нарушителей кибербезопасности при реализации целевых атак // Вопросы кибербезопасности. 2022. № 4(50). С. 52–79. DOI: 10.21681/2311-3456-2022-4-52-79.
28. Федорченко Е. В., Котенко И. В., Федорченко А. В., Новикова Е. С., Саенко И. Б. Оценивание защищенности информационных систем на основе графовой модели эксплойтов // Вопросы кибербезопасности. 2023. № 3(57). С. 23–36. DOI: 10.21681/2311-3456-2023-3-23-36.
29. Васильев В. И., Вульфин А. М., Картак В. М., Башмаков Н. М., Кириллова А. Д. Распределенная система обнаружения сетевых атак на основе федеративного трансферного обучения // Вопросы кибербезопасности. 2024. № 6. С. 117–129. DOI: 10.21681/2311-3456-2024-6-117-129.

30. Tushkanova O., Levshun D., Branitskiy A., Fedorchenko E., Novikova E., Kotenko I. Detection of Cyberattacks and Anomalies in Cyber-Physical Systems: Approaches, Data Sources, Evaluation // Algorithms. 2023. Vol. 16, No. 2. Article 85. DOI: 10.3390/a16020085.
31. Котенко И. В., Дун Х. Обнаружение атак в Интернете вещей на основе многозадачного обучения и гибридных методов сэмплирования // Вопросы кибербезопасности. 2024. № 2(60). С. 10–21. DOI: 10.21681/2311-3456-2024-2-10-21.
32. Левшун Д. С., Веснин Д. В., Котенко И. В. Прогнозирование категорий уязвимостей в конфигурациях устройств с помощью методов искусственного интеллекта // Вопросы кибербезопасности. 2024. № 3(61). С. 33–39. DOI: 10.21681/2311-3456-2024-3-33-39.

EXPLAINABLE INTERPRETATION OF INCIDENTS BASED ON A LARGE LANGUAGE MODEL AND A RETRIEVAL-AUGMENTED GENERATION

Kotenko I. V.[7], *Abramenko G. T.*[8]

Keywords: *: cybersecurity; intrusion detection and prevention systems; threat hunting; explainable artificial intelligence; knowledge graph; deep learning: heterogeneous graph neural network; MITRE ATT&CK; LLM; RAG; Suricata.*

The purpose of the study: *to increase the reliability and explainability of interpreting Suricata IDS alerts by means of an ontologically enriched knowledge graph, heterogeneous graph representations, and Retrieval-Augmented Generation (RAG) based on a local large language model.*

Research methods: *the construction of an ontology-guided knowledge graph that links Suricata data to MITRE ATT&CK tactics/techniques is the first stage of the research. Following this, a heterogeneous graph neural network (HGNN) is trained to obtain contextual node embeddings. The retrieval of relevant context via nearest neighbours in the embedding space is then conducted. Finally, a local LLM (about 7B parameters) is generated within a RAG pipeline to generate explanations. An experimental evaluation on a corpus of approximately 25,000 Suricata alerts is then conducted using metrics of interpretation accuracy, hallucination rate, and relevance.*

Results obtained: *an ontology-guided method for interpreting Suricata alerts was developed, providing more complete and accurate explanations than a baseline approach. The utilisation of ontologies has been demonstrated to enhance the substantive content of explanations by approximately 15%, while concomitantly resulting in only a marginal increase in generation time. The integration of ontology and a heterogeneous knowledge graph has been demonstrated to improve the correct mapping of alerts to MITRE ATT&CK techniques and reduce the risk of erroneous explanations.*

Scientific novelty: *the integration of an ontology and a heterogeneous knowledge graph with RAG-based generation over a local LLM to anchor low-level IDPS events to MITRE ATT&CK techniques has been suggested; the applicability of ontologies to explainable AI in cybersecurity has been shown.*

References

1. Singh A. Contextual Threat Intelligence and Alert Prioritization with Foundation-Sec-8B // International Journal of Artificial Intelligence Research and Development. 2025. Vol. 3, No. 1. P. 131–145. DOI: 10.34218/IJAIRD_03_01_009.
2. Arreche O., Guntur T., Abdallah M. XAI-IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Enhancing Network Intrusion Detection Systems // Applied Sciences. 2024. Vol. 14, No. 10. Article 4170. DOI: 10.3390/app14104170.
3. Hassanin M., Moustafa N. A Comprehensive Overview of Large Language Models (LLMs) for Cyber Defences: Opportunities and Directions // arXiv preprint arXiv:2405.14487. 2024. DOI: 10.48550/arXiv.2405.14487.
4. Fan W., Ding Y., Ning L., Wang S., Li H., Yin D., Chua T.-S., Li Q. A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models // Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'24). 2024. P. 6491–6501.
5. Mavromatis C., Karypis G. GNN-RAG: Graph Neural Retrieval for Large Language Model Reasoning // arXiv preprint arXiv:2405.20139. 2024. DOI: 10.48550/arXiv.2405.20139.
6. Wu D., Yan Y., Liu Z., Liu Z., Sun M. KG-Infused RAG: Augmenting Corpus-Based RAG with External Knowledge Graphs // arXiv preprint arXiv:2506.09542. 2025. DOI: 10.48550/arXiv.2506.09542.
7. Al-Sada B., Sadighian A., Oligeri G. MITRE ATT&CK: State of the Art and Way Forward // ACM Computing Surveys. 2024. Vol. 57, No. 1. Article 12. P. 1–37. DOI: 10.1145/3687300.
8. Li H., Shi Z., Pan C., Zhao D., Sun N. Cybersecurity Knowledge Graphs Construction and Quality Assessment // Complex & Intelligent Systems. 2024. Vol. 10. P. 1201–1217. DOI: 10.1007/s40747-023-01205-1.
9. Hemberg E., Kelly J., Shlapentokh-Rothman M., Reinstadler B., Xu K., Rutar N., O'Reilly U.-M. Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Platforms for Cyber Hunting // arXiv preprint arXiv:2010.00533. 2021. DOI: 10.48550/arXiv.2010.00533.
10. Freitas S., Gharib A. GraphWeaver: Billion-Scale Cybersecurity Incident Correlation // Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM'24). 2024. P. 4479–4486. DOI: 10.1145/3627673.3680057.

7 Igor V. Kotenko, Honored Worker of Science of the Russian Federation, Dr.Sc. of Technical Sciences, Professor, Chief Scientist and Head of Laboratory of Computer Security Problems at St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg, Russia. E-mail: ivkote@comsec.spb.ru
8 Georgii T. Abramenko, Ph.D. Student, Faculty of Secure Information Technologies, ITMO University (National Research University ITMO), St Petersburg, Russia. E-mail: gtabramenko@itmo.ru

11. Sikos L. F. Cybersecurity Knowledge Graphs // Knowledge and Information Systems. 2023. Vol. 65, No. 9. P. 3511–3531. DOI: 10.1007/s10115-023-01860-3.
12. Zhao X., Jiang R., Han Y., Li A., Peng Z. A Survey on Cybersecurity Knowledge Graph Construction // Computers & Security. 2024. Vol. 136. Article 103524. DOI: 10.1016/j.cose.2023.103524.
13. Xiong W., Legrand E., Åberg O., Lagerström R. Cyber Security Threat Modeling Based on the MITRE Enterprise ATT&CK Matrix // Software and Systems Modeling. 2022. Vol. 21. P. 157–177. DOI: 10.1007/s10270-021-00898-7.
14. Arikkat D. R., Abhinav M., Binu N., Rajendran N., Bhattacharjee D., Das R. K., Ajay C. R., Shiyam Sundar R., Bhavesh S. IntelIBot: A Retrieval-Augmented Large Language Model Chatbot for Cyber Threat Knowledge Delivery // arXiv preprint arXiv:2411.05442. 2024. DOI: 10.48550/arXiv.2411.05442.
15. Kurniawan K., Kiesling E., Ekelhart A. CyKG-RAG: Towards Knowledge-Graph Enhanced Retrieval-Augmented Generation for Cybersecurity // CEUR Workshop Proceedings. 2024. Vol. 3950. P. 51–64.
16. Tellache A., Amara-Korba A., Mokhtari A., Moldovan H., Ghamri-Doudane Y. Advancing Autonomous Incident Response: Leveraging LLMs and Cyber Threat Intelligence // arXiv preprint arXiv:2508.10677. 2025. DOI: 10.48550/arXiv.2508.10677.
17. Qi Y., Gu Z., Li A., Zhang X., Shafiq M., Mei Y., Lin K. Cybersecurity Knowledge Graph Enabled Attack Chain Detection for Cyber-Physical Systems // Computers & Electrical Engineering. 2023. Vol. 108. Art. 108660. DOI: 10.1016/j.compeleceng.2023.108660.
18. Ekelhart A., Ekaputra F.J., Kiesling E. SLOGERT: Automated Log Knowledge Graph Construction // The Semantic Web – ESWC 2021. Lecture Notes in Computer Science. 2021. Vol. 12731. P. 219–234. DOI: 10.1007/978-3-030-77385-4_16.
19. Kurniawan K., Ekelhart A., Kiesling E., Winkler D., Quirchmayr G., Tjoa A.M. Virtual Knowledge Graphs for Federated Log Analysis // Proceedings of the 16th International Conference on Availability, Reliability and Security (ARES'21). 2021. Article 50. 10 p. DOI: 10.1145/3465481.3470077.
20. Bilot T., El-Madhoun N., Al-Agha K., Zouaoui A. Graph Neural Networks for Intrusion Detection: A Survey // IEEE Access. 2023. Vol. 11. P. 49114–49139. DOI: 10.1109/ACCESS.2023.3275789.
21. Zhong M., Lin M., Zhang C., Xu Z. A Survey on Graph Neural Networks for Intrusion Detection Systems: Methods, Trends and Challenges // Computers & Security. 2024. Vol. 141. Article 103821. DOI: 10.1016/j.cose.2024.103821.
22. Fang J., Liu W., Gao Y., Liu Z., Zhang A., Wang X., He X. Evaluating Post-hoc Explanations for Graph Neural Networks via Robustness Analysis // Advances in Neural Information Processing Systems (NeurIPS'23), Oral. 2023.
23. Lewis P., Perez E., Piktus A., Petroni F., Karpukhin V., Goyal N., Küttler H., Lewis M., Yih W.-t., Rocktäschel T., Riedel S., Kiela D. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // Advances in Neural Information Processing Systems. 2020. Vol. 33. P. 9459–9474. DOI: 10.5555/3495724.3495881.
24. Zhu X., Xie Y., Liu Y., Li Y., Hu W. Knowledge Graph-Guided Retrieval-Augmented Generation (KG²RAG) // Proceedings of NAACL 2025 (Long Papers). 2025. P. 8912–8924. DOI: 10.18653/v1/2025.naacl-long.449.
25. Farrukh Y. A., Wali S., Khan I., Bastian N. D. Xg-nid: Dual-modality network intrusion detection using a heterogeneous graph neural network and large language model // Expert Systems with Applications. 2025. p. 128089.
26. Novikova, E. S.; Bukhtiyarov, M. A.; Kotenko, I. V.; Saenko, I. B.; Fedorchenko, E. V. Intrusion Detection Based on Federated Learning: System Architecture and Experiments. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2023, no. 6(58), 50–66. https://doi.org/10.21681/2311-3456-2023-6-50-66.
27. Kotenko, I. V.; Khmyrov, S. S. Analysis of Models and Methods Used for Attribution of Cybersecurity Adversaries in Targeted Attacks. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2022, no. 4(50), 52–79. https://doi.org/10.21681/2311-3456-2022-4-52-79.
28. Fedorchenko, E. V.; Kotenko, I. V.; Fedorchenko, A. V.; Novikova, E. S.; Saenko, I.B. Assessing the Security of Information Systems Based on a Graph Model of Exploits. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2023, no. 3(57), 23–36. https://doi.org/10.21681/2311-3456-2023-3-23-36.
29. Vasiliev, V. I.; Vulfin, A. M.; Kartak, V. M.; Bashmakov, N. M.; Kirillova, A. D. Distributed Network Attack Detection System Based on Federated Transfer Learning. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 6, 117–129. https://doi.org/10.21681/2311-3456-2024-6-117-129.
30. Tushkanova, O.; Levshun, D.; Branitskiy, A.; Fedorchenko, E.; Novikova, E.; Kotenko, I. Detection of Cyberattacks and Anomalies in Cyber-Physical Systems: Approaches, Data Sources, Evaluation. Algorithms, 2023, 16(2), 85. https://doi.org/10.3390/a16020085.
31. Kotenko, I. V.; Dun, H. Detection of Attacks in the Internet of Things Based on Multitask Learning and Hybrid Sampling Methods. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 2(60), 10–21. https://doi.org/10.21681/2311-3456-2024-2-10-21.
32. Levshun, D. S.; Vesnin, D. V.; Kotenko, I. V. Predicting Categories of Vulnerabilities in Device Configurations Using Artificial Intelligence Methods. Voprosy kiberbezopasnosti [Cybersecurity Issues], 2024, no. 3(61), 33–39. https://doi.org/10.21681/2311-3456-2024-3-33-39.