# ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ФОРМИРОВАНИЯ ДОКАЗАТЕЛЬНОЙ БАЗЫ ПРЕСТУПЛЕНИЙ В УМНЫХ ДОМАХ

***Абдурахимов Бахтиёр Файзиевич***

***Преподаватель Национального Университета Узбекистана, город Ташкент, Республика Узбекистан***

***Алланов Ориф Менглимуратович***

***Заведующий Кафедры Кибербезопасности и Криминалистики Ташкентского Университета Информационных технологий, город Ташкент, Республика Узбекистан***

***Юлдашева Нафиса Салимовна***

***Преподаватель Ташкентского Университета Информационных технологий, город Ташкент, Республика Узбекистан***

***Янгазова-Курмаева Маликахон Рустамовна***

***Студент Магистратуры Ташкентского Университета Информационных технологий, город Ташкент, Республика Узбекистан***

***Аннотация:*** *В работе описывается решение задачи формирования доказательной базы преступлений, совершенных с использованием IoT-устройств в умных домах путем использования искусственного интеллекта в качестве помощника в определении возможных доказательств правонарушений. Результаты анализа представляются в процентах, отражающих вероятность, что те или иные события могут быть связаны с правонарушением.*

***Ключевые слова:*** *аномалия, доказательство, интернет вещей (IoT), искусственный интеллект (AI), модель, преступление, умный дом, экосистема.*

# APPLICATION OF ARTIFICIAL INTELLIGENCE FOR BUILDING THE EVIDENCE BASE OF CRIMES IN SMART HOMES

***Abdurakhimov Bakhtiyor Fayzievich***

***Lecturer, National University of Uzbekistan, Tashkent, Republic of Uzbekistan***

***Allanov Orif Menglimuratovich***

***Head of the Department of Cybersecurity and Forensics, Tashkent University of Information Technologies, Tashkent, Republic of Uzbekistan***
***Yuldasheva Nafisa Salimovna***
***Lecturer, National University of Uzbekistan, Tashkent, Republic of Uzbekistan***
***Yangazova-Kurmaeva Malikakhon Rustamovna***
***Master's Student, Tashkent University of Information Technologies, Tashkent, Republic of Uzbekistan***

***Abstract:*** *This paper describes a solution to the problem of forming an evidence base for crimes committed using IoT devices in smart homes by employing artificial intelligence as an assistant in identifying potential evidence of offenses. The results of the analysis are presented as percentages, reflecting the probability that certain events may be related to a criminal act.*

***Keywords:*** *anomaly, evidence, Internet of Things (IoT), artificial intelligence (AI), model, crime, smart home, ecosystem.*

Умный дом — это концепция жилого пространства, оснащённого интеллектуальными технологиями, которые обеспечивают автоматическое и/или дистанционное управление различными системами и устройствами. Такие дома используют интегрированные системы управления для повышения комфорта, энергоэффективности, безопасности и удобства проживания. Умный дом является одним из ярких примеров применения интернета вещей (IoT). В умном доме используется множество устройств, которые подключены к сети и взаимодействуют друг с другом. Интернет вещей (IoT, Internet of Things) — это концепция, которая описывает сеть физически подключённых устройств (вещей), способных собирать, обмениваться и обрабатывать данные через интернет. Эти устройства оснащены датчиками, процессорами, программным обеспечением и коммуникационными модулями, что позволяет им взаимодействовать с другими устройствами и системами без участия человека. IoT-устройства, такие как системы видеонаблюдения, умные замки и датчики, постоянно генерируют огромные объемы данных, которые могут стать важными

доказательствами в случае преступлений. Однако стандартные методы обработки и анализа данных не всегда подходят для эффективного выявления потенциальных правонарушений в реальном времени. В связи с этим актуальной задачей становится разработка решений на базе искусственного интеллекта, которые позволят анализировать данные IoT-систем для создания доказательной базы, особенно в условиях сложных экосистем умных домов.

На сегодняшний день исследователи активно изучают вопросы безопасности IoT и умных домов, включая методы аномалий, диагностики и анализа данных. Множество научных работ сосредоточено на детектировании угроз безопасности в IoT-средах, а также на использовании искусственного интеллекта и машинного обучения для обработки больших данных. Тем не менее, исследования, направленные на формирование доказательной базы с использованием AI, остаются относительно новыми и актуальными. Значительное внимание уделяется детектированию аномалий, но его внедрение в области криминалистики IoT-устройств только начинает развиваться. Так, Yaman Salem, Amani Yousef Owda в своей работе предложили многоуровневую систему, которая охватывает не только сбор, но и валидацию, а также корреляцию данных между устройствами, чтобы установить последовательность событий в рамках одной преступной деятельности. Основной акцент сделан на автоматизации и стандартизации процессов для IoT-экспертизы, что позволяет сделать судебные экспертизы более целостными и надежными. Abiodun Abdullahi Solanke описывает, как технологии ИИ могут способствовать улучшению криминалистического анализа цифровых данных. Ямана Салема, Маджи Оды и Амании Юсеф Оды, посвящено разработке многоуровневой цифрового судебно-экспертного фреймворка для устройств Интернета вещей (IoT). Ключевая особенность — её способность охватывать все уровни архитектуры IoT и фокусироваться на артефактах интереса (AoI). В исследовании вводится матрица действия/обнаружения (Action/Detection Matrix), которая помогает систематизировать и ускорить процесс судебной экспертизы в IoT-средах. Проведённые эксперименты показали, что предложенная рамка

превосходит существующие методы по таким параметрам, как удобство использования, полнота охвата, фокус на AoI и ускорение процесса расследования. Таким образом, исследование представляет собой значимый вклад в область цифровой судебной экспертизы, предлагая адаптированную и эффективную методологию для анализа данных, генерируемых IoT-устройствами.

Однако эти работы рассматривают проблему на этапе получения данных из датчиков устройств интернета вещей до их обработки центральным контроллером системы. Более того, хоть и использование искусственного интеллекта затронуто как способ добычи доказательств с помощью корреляции данных, в них отсутствует разработка модели.

В работе представлена методология использования ИИ в выявлении доказательств преступлений умных домов, а также представлена модель ИИ, выявляющая потенциальные доказательства из имеющихся данных и выдающая процент вероятности отношения к преступлению аномалий и коррелированных данных. Для анализа и выявления доказательств используются структурированные данные из лидирующих экосистем умных домов на рынке: Google Home, Apple Homekit и Amazon Alexa.

Объектом исследования является процесс формирования доказательной базы преступлений, совершаемых с использованием IoT-устройств в умных домах.

Предметом исследования выступает использование моделей искусственного интеллекта для анализа данных экосистем умных домов, с целью выявления аномальных действий и последовательностей событий.

Экосистема умного дома — это сеть устройств и технологий, которые работают вместе для автоматизации и упрощения управления домом.

Центральной частью умного дома часто является центральный контроллер или приложение, которое связывает все устройства и позволяет управлять ими удаленно через смартфон или голосовые команды с помощью ассистентов. В данном исследовании рассматриваются 3 лидирующие по количеству пользователей экосистемы: Google Home, Apple Homekit и Amazon Alexa.

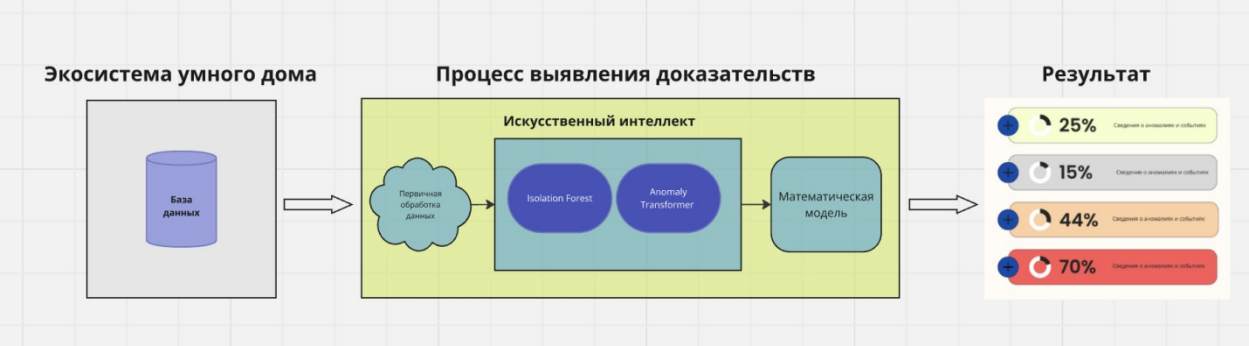
1. Google Home: управляется через Google Assistant и идеально подходит для пользователей экосистемы Google.
2. Apple HomeKit: это система, ориентированная на пользователей устройств Apple. Она предлагает высокую безопасность и автоматизацию, а также отличную интеграцию с продуктами Apple, такими как iPhone, iPad, Apple TV и HomePod.
3. Amazon Alexa: основана на голосовом помощнике Amazon Alexa и предлагает обширную совместимость с различными умными устройствами от разных производителей.

Каждая из трех систем имеет 100 и более миллионов пользователей.

Для эффективного формирования доказательной базы преступлений в системах Интернета вещей (IoT) умных домов с использованием искусственного интеллекта (ИИ), можно выделить следующие ключевые этапы:

1- Рисунок

Методология использования искусственного интеллекта для выявления доказательств




1. Извлечение данных. Необходимо получить доступ к данным устройств из хранилищ, включая журналы событий, показания датчиков и данные об обычном поведении пользователя. Для этого будут использоваться следующие таблицы базы данных экосистемы умного дома:
   - События
   - Устройства

- Расписание

2. Обработка данных для передачи на анализ ИИ. Следующим шагом полученные данные должны быть очищены и нормализованы, т. е. без пропусков, выбросов и несоответствий. Также объединение данных из различных таблиц в единую структуру, обеспечив их совместимость для дальнейшего анализа осуществляется на данном этапе.
3. Анализ данных с использованием ИИ. Полученные и подготовленные данные отправляются в искусственный интеллект на анализ. В данном решение нами предложено использование двух элементов анализа:

- Isolation Forest – алгоритм для выявления аномалий по отношению к обычному поведению пользователя. Аномальные точки требуют меньшего количества разделений для изоляции, что позволяет алгоритму эффективно их обнаруживать. Алгоритм основывается на концепции изоляции аномальных точек в данных. Его основная идея заключается в том, что аномалии легче изолировать, чем нормальные точки, так как они встречаются реже и отличаются от общей массы данных. Алгоритм базируется на следующих ключевых принципах:
  - Количество разбиений для изоляции: Аномальные точки требуют меньшего количества разбиений (splits) в случайно построенном дереве для изоляции, в сравнении с обычными точками.
  - Оценка аномальности (Anomaly Score): Для определения степени аномальности точки используется средняя длина пути (h(x)), которая вычисляется для множества деревьев, и поправочный коэффициент нормализации (c(n)), относящийся к количеству точек в выборке.
  - . Чем короче путь изоляции, тем выше вероятность того, что точка является аномальной.

Показатель аномалии, представленный коэффициентом нормализации (s(x,n)) относительно размера выборки (n), рассчитывается как:

$$s(x,n)=2^{\frac{-h(x)}{c(n)}}$$

Если s(x,n) близко к 1, то точка x считается аномальной. Если s(x,n) близко к 0.5, то точка скорее всего является нормальной.

- Anomaly Transformer в unsupervised режиме – для выявления аномалий в последовательностях данных. Anomaly Transformer — это алгоритм машинного обучения, предназначенный для обнаружения аномалий в временных рядах в несупервизорном режиме. Он фокусируется на различиях в ассоциациях между нормальными и аномальными точками во временном ряду, что позволяет эффективно выявлять отклонения. Внутри трансформера используется механизм анализ разрывов внимания (Anomaly-Attention Discrepancy Analysis), который сопоставляет степень важности каждой точки временного ряда с её соседями. Основная идея: если внимание, выделяемое на текущую точку, значительно отличается от ожиданий (по сравнению с её "нормальными" соседями), то эта точка может быть аномальной. Anomaly Transformer использует два ключевых компонента:
  - Карта внимания (A) Представляет собой матрицу, показывающую, насколько каждый временной шаг (или элемент входного ряда) связан с остальными временными шагами. Каждый элемент (A(i,j)) вычисляется с использованием механизма внимания масштабированного скалярного произведения, где сходство между вектором запроса на временном шаге ($Q_i$) и ключевым вектором на временном шаге ($K_j$¿) определяется их скалярным произведением ($Q_i K_j^T$). Этот показатель сходства масштабируется по квадратному корню ключевой размерности ($d_k$) для поддержания числовой стабильности, а затем передается через функцию softmax, чтобы обеспечить правильное взвешивание на всех временных шагах.

$$A(i,j) = Softmax\left(\frac{Q_i K_j^T}{\sqrt{d_k}}\right)$$

  - Loss-функция (анализ разрывов внимания): Loss-функция делится на два компонента:
    1) $L_{recon}$: для минимизации ошибки восстановления временного ряда.
    2) $L_{prior}$: для измерения расхождения между распределением внимания текущей точки и ожидаемым распределением.

Общая функция потерь с балансирующим коэффициентом (λ):

$$L = L_{recon} + ¿\lambda L_{prior},$$

где λ — коэффициент для баланса двух частей.

4. Результатами обоих алгоритмов являются оценки аномальности (anomaly_scores), которые присваиваются каждой точке временного ряда на основе её степени аномальности. Эти оценки нормализуются (normal_scores) так, чтобы они попадали в диапазон [0, 1] с использованием минимального (min_score) и максимального (max_score) возможных оценок:

$$normal_{scores} = \frac{anomaly_{scores} - min_{score}}{max_{score} - min_{score}}$$

5. После, результаты каждой точки (normal_scores) конвертируются в процентную вероятность (probabilities_percent):

$$probabilities_{percent} = normal_{scores} \times 100$$

Результатом работы алгоритмов (P) является полу сумма вероятности аномалии относительно обычного поведения пользователя ($P_e$) и последовательной вероятности для каждой точки ($P_s$).

$$P = \frac{P_e + P_s}{2}$$

6. Финальное решение принимается криминалистом.

1) На основе представленных процентных вероятностей и предоставленных данных специалист опирается на цифровой показатель: если он ниже определенного значение заданного самим специалистом и не имеет явного отношения к составу преступления, то данный результат не принимается во внимание.
2) Однако, даже если результат вероятности высок, но приведенный список аномалий и событий не является достаточно полным, то криминалист может отклонить данный результат или же провести дополнительный анализ данных.

3) Если же показатель стремится к плану поставленной специалистом и имеет полный объем необходимых коррелированных сведений, то этот результат определенно принимается во внимание и формирует доказательство.

## ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Проведено исследование по работе искусственного интеллекта и экспертной модели. Эксперимент прошел по следующим этапам:

1. Имитирование работы экосистемы умного дома. В первую очередь была создана и заполнена база данных, эмитирующая базу данных экосистемы умного дома. Были созданы таблицы событий (events), пользователей (users), устройств(devices) и расписаний(schedules).
2. После заполнения базы данных начинается работа искусственного интеллекта – последовательно отрабатывают два алгоритма: Isolation Forest и Anomaly Transformer.

- Результатами работы Isolation Forest является процентная вероятность отношения/аномалии события относительно обычного поведения пользователя, представленного в таблице расписаний. Результаты состоят из порядкового номера события в базе, действия события и его вероятности.

2- Рисунок

Результаты работы Isolation Forest

```
ai_digital_forensics    | Аномальные события (Isolation Forest):
ai_digital_forensics    | {'id': 1, 'event': 'Light turned on', 'event_probability': '93.55%'}
ai_digital_forensics    | {'id': 2, 'event': 'Speaker played music', 'event_probability': '0.68%'}
ai_digital_forensics    | {'id': 3, 'event': 'Speaker played music', 'event_probability': '0.68%'}
ai_digital_forensics    | {'id': 4, 'event': 'Speaker played music', 'event_probability': '0.68%'}
ai_digital_forensics    | {'id': 5, 'event': 'Camera activated', 'event_probability': '0.0%'}
ai_digital_forensics    | {'id': 6, 'event': 'Lock engaged', 'event_probability': '30.86%'}
ai_digital_forensics    | {'id': 7, 'event': 'Battery low', 'event_probability': '56.47%'}
ai_digital_forensics    | {'id': 8, 'event': 'Lock engaged', 'event_probability': '30.86%'}
ai_digital_forensics    | {'id': 9, 'event': 'Light turned on', 'event_probability': '93.55%'}
```

- Результатами работы Anomaly Transformer является процентная вероятность отношения/аномалии последовательности событий относительно других событий. События разделяются на

последовательности длинной 21 элемент и отображаются в виде порядкового номера последовательности и ее вероятности.

3- Рисунок

Результаты работы Anomaly Transformer

```
ai_digital_forensics     | Аномальные последовательности (Anomaly Transformer):
ai_digital_forensics     | {'sequence_index': 0, 'sequence_probability': '0.0%'}
ai_digital_forensics     | {'sequence_index': 1, 'sequence_probability': '100.0%'}
```

3. Финальным этапом эксперимента являлась обработка результатов математической моделью. Для этого каждое событие было оценено в частности: по порядковому номеру события определялось его отношение в определенной последовательности и ее вероятность. Получив последовательную вероятность, математическая модель подсчитывает итоговый результат для каждой точки.

4- Рисунок

Результаты математической модели

```
ai_digital_forensics     | Объединенные результаты (события + последовательности):
ai_digital_forensics     | {'id': 1, 'event': 'Light turned on', 'event_probability': '93.55%', 'sequence_probability': '0.0%', 'combined_probability': '46.77%'}
ai_digital_forensics     | {'id': 2, 'event': 'Speaker played music', 'event_probability': '0.68%', 'sequence_probability': '0.0%', 'combined_probability': '0.34%'}
ai_digital_forensics     | {'id': 3, 'event': 'Speaker played music', 'event_probability': '0.68%', 'sequence_probability': '0.0%', 'combined_probability': '0.34%'}
ai_digital_forensics     | {'id': 4, 'event': 'Speaker played music', 'event_probability': '0.68%', 'sequence_probability': '0.0%', 'combined_probability': '0.34%'}
ai_digital_forensics     | {'id': 5, 'event': 'Camera activated', 'event_probability': '0.0%', 'sequence_probability': '0.0%', 'combined_probability': '0.0%'}
ai_digital_forensics     | {'id': 6, 'event': 'Lock engaged', 'event_probability': '30.86%', 'sequence_probability': '0.0%', 'combined_probability': '15.43%'}
ai_digital_forensics     | {'id': 7, 'event': 'Battery low', 'event_probability': '56.47%', 'sequence_probability': '0.0%', 'combined_probability': '28.23%'}
ai_digital_forensics     | {'id': 8, 'event': 'Lock engaged', 'event_probability': '30.86%', 'sequence_probability': '0.0%', 'combined_probability': '15.43%'}
```

В работе предложена методология использования искусственного интеллекта (ИИ) для формирования доказательной базы преступлений, совершённых с использованием IoT-устройств в умных домах. Разработанная система включает этапы получения данных из баз данных экосистем умного дома, их предварительной обработки, анализа с применением ИИ и представления результатов в удобной для специалистов форме. Применение ИИ позволяет эффективно выявлять аномальные действия и последовательности событий, указывающие на потенциальные правонарушения, что способствует повышению качества и скорости расследований в сфере кибербезопасности.

Новизна работы заключается в разработке интегрированной методологии, объединяющей искусственный интеллект (ИИ) и экспертной оценки для анализа данных из экосистем умных домов с целью формирования доказательной базы преступлений, совершённых с использованием IoT-устройств.

В отличие от существующих исследований, которые преимущественно фокусируются на отдельных аспектах, таких как сбор данных или детектирование аномалий, представленная методология охватывает полный цикл: от получения данных из баз данных умных домов до их анализа с помощью ИИ и представления результатов в удобной для специалистов форме. Это позволяет более эффективно выявлять и интерпретировать аномальные действия и последовательности событий, потенциально связанные с правонарушениями, что ранее не было полноценно реализовано в подобных системах.

Предлагаемая экспетная (математическая) модель позволяет количественно оценивать вероятность того, что определённые события, зафиксированные в экосистемах умных домов, связаны с преступной деятельностью. Такая модель обеспечивает объективность и точность в процессе формирования доказательной базы, что является существенным вкладом в область цифровой криминалистики.

**Список Литературы:**


1. Эштон К. Та самая «вещь» под названием Интернет вещей // RFID Journal. 2009.
2. Губби Дж., Буйя Р., Марушич С., Паланисвами М. Интернет вещей (IoT): концепция, архитектурные элементы и направления развития // Future Generation Computer Systems. 2013. Т. 29, № 7. С. 1645–1660.
3. Дзанелла А., Буй Н., Кастеллани А., Вангелиста М., Дзорци М. Интернет вещей для умных городов // IEEE Internet of Things Journal. 2014. Т. 1, № 1. С. 22–32.
4. Миора́нди Д., Сикари С., Де Пеллегрини Ф., Хламтак И. Интернет вещей:

концепция, приложения и исследовательские задачи // Ad Hoc Networks. 2012. Т. 10, № 7. С. 1497–1516.

5. Манн Д., Фролоу М. Умные дома и Интернет вещей: возможности, тенденции и проблемы. Берлин: Springer, 2015.
6. Саламх Ф. Е. Модель криминалистического анализа домашних автоматизированных устройств (FAHAD): кейс Kasa Smart Light Bulb и Eufy Floodlight Camera // International Journal of Cyber Forensics and Advanced Threat Investigations. 2020.
7. Соланке А. А. Цифровая криминалистика и искусственный интеллект: оценка, стандартизация и оптимизация методов извлечения цифровых доказательств // KI – Künstliche Intelligenz. 2022.
8. Салем Я., Овда А. Ю. К цифровой криминалистике 4.0: многоуровневая рамочная модель цифровой криминалистики для устройств Интернета вещей // International Journal of Wireless and Microwave Technologies. 2024.
9. Кафле К., Моран К., Мандхар С., Надкарни А., Пошиваник Д. Исследование систем хранения данных в домашней автоматизации // Proceedings of the Ninth ACM Conference on Data and Application Security and Privacy. 2019.
10. Лю Ф. Т., Тинг К. М., Чжоу Ч. Х. Isolation Forest // Proceedings of the 2008 IEEE International Conference on Data Mining (ICDM). 2008.
11. Чандола В., Банерджи А., Кумар В. Обзор методов обнаружения аномалий // ACM Computing Surveys. 2009. Т. 41, № 3.
12. Филонов П. и др. Принципы и алгоритмы обнаружения аномалий. ISBN 978-1733455934. 2020.
13. Зонг Б., Сонг Л., Е Ц. Глубокая автоэнкодерная гауссовская смесь для неучетного обнаружения аномалий // Proceedings of the 2018 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2018. С. 118–127.
14. Дай С., Хэ С. Обнаружение аномалий с помощью LSTM-автоэнкодеров // 2018 IEEE International Conference on Big Data (Big Data). 2018. С. 1485–1494.

15. Ли С., Чжан Л. Transformer для обнаружения аномалий: архитектура на основе Transformer для временных рядов // Proceedings of the 2020 Conference on Neural Information Processing Systems (NeurIPS). 2020.
16. Ахмед М., Махмуд А. Н., Ху Д. Обзор методов обнаружения сетевых аномалий // Journal of Network and Computer Applications. 2016. Т. 60. С. 19–31.
17. Бергман П., Фаузер Б., Саттлеггер Д., Зик Б. Внутреннее обнаружение выбросов с помощью LSTM для одно-классовой классификации // Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019. С. 6557–6566.
18. Ся Ф., Лю С. Обзор методов обнаружения аномалий во временных рядах // IEEE Access. 2021. Т. 9. С. 22258–22272.
19. Чжао Ц., Лю В. Модели машинного обучения для обнаружения преступлений в умных домах // Journal of Cyber Security and Privacy. 2020.