Аннотация – В статье "MediHunt: A Network Forensics Framework for Medical IoT Devices" представлена разработка MediHunt, автоматической платформы сетевой криминалистики, предназначенной для обнаружения атак сетевого трафика на основе потоков в сетях MQTT, которые обычно используются в средах интеллектуальных больниц. MediHunt может обнаруживать различные атаки уровня TCP/IP и уровня приложений в сетях MQTT, используя модели машинного обучения. Платформа направлена на расширение возможностей криминалистического анализа в средах MIoT, обеспечивая эффективное отслеживание вредоносных действий и смягчение их последствий.

I. Введение

В документе "MediHunt: A Network Forensics Framework for Medical IoT Devices" рассматривается необходимость надёжной сетевой криминалистики в медицинских средах Интернета вещей (MIoT), особенно с упором на сети MQTT. Эти сети обычно используются в интеллектуальных больничных средах благодаря их облегчённому протоколу связи. Освещаются проблемы обеспечения безопасности устройств MIoT, которые часто ограничены в ресурсах и обладают ограниченной вычислительной мощностью. В качестве серьёзной проблемы упоминается отсутствие общедоступных потоковых наборов данных, специфичных для MQTT, для обучения систем обнаружения атак.

В документе представлен MediHunt как решение для автоматизированной сетевой криминалистики, предназначенное для обнаружения атак на основе сетевого трафика в сетях MQTT в режиме реального времени. Его цель – предоставить комплексное решение для сбора данных, анализа, обнаружения атак, представления и сохранения доказательств. Он разработан для обнаружения различных уровней TCP / IP и атак прикладного уровня в сетях MQTT и использует модели машинного обучения для расширения возможностей обнаружения и подходит для развёртывания на устройствах MIoT с ограниченными ресурсами.

II. Преимущества и недостатки предлагаемого решения

A. *Преимущества*

- **Обнаружение атак в режиме реального времени**: MediHunt предназначен для обнаружения атак на основе сетевого трафика в режиме реального времени для уменьшения потенциального ущерба и обеспечения безопасности сред MIoT.
- **Комплексные возможности криминалистики**: Платформа предоставляет комплексное решение для сбора данных, анализа, обнаружения атак, представления и сохранения доказательств. Это делает его надёжным инструментом сетевой криминалистики в средах MIoT.
- **Интеграция с машинным обучением**: Используя модели машинного обучения, MediHunt расширяет свои возможности обнаружения. Использование пользовательского набора данных, который включает данные о потоках как для атак уровня TCP/IP, так и для атак прикладного уровня, позволяет более точно и эффективно обнаруживать широкий спектр кибератак.
- **Высокая производительность**: решение показало высокую производительность, получив баллы F1 и точность обнаружения, превышающую 0,99 и указывает на то, что она обладает высокой надёжностью при обнаружении атак на сети MQTT.
- **Эффективность использования ресурсов**: несмотря на свои широкие возможности, MediHunt разработан с учётом экономии ресурсов, что делает его подходящим для развёртывания на устройствах MIoT с ограниченными ресурсами (raspberry Pi).

B. *Недостатки*

- **Ограничения набора данных**: хотя MediHunt использует пользовательский набор данных для обучения своих моделей машинного обучения, создание и обслуживание таких наборов данных может быть сложной задачей. Набор данных необходимо регулярно обновлять, чтобы охватывать новые и зарождающиеся сценарии атак.
- **Ограничения ресурсов**: хотя MediHunt разработан с учётом экономии ресурсов, ограничения, присущие устройствам MIoT, такие как ограниченная вычислительная мощность и память, все ещё могут создавать проблемы. Обеспечить бесперебойную работу фреймворка на этих устройствах без ущерба для их основных функций может быть непросто.
- **Сложность реализации**: Внедрение и поддержка платформы сетевой криминалистики на основе машинного обучения может быть сложной задачей. Это требует опыта в области кибербезопасности и машинного обучения, который может быть доступен не во всех медицинских учреждениях.
- **Зависимость от моделей машинного обучения**: Эффективность MediHunt в значительной степени зависит от точности и надёжности его моделей машинного обучения. Эти модели необходимо

обучать на высококачественных данных и регулярно обновлять, чтобы они оставались эффективными против новых типов атак.

- **Проблемы с масштабируемостью**: хотя платформа подходит для небольших развёртываний на устройствах типа Raspberry Pi, ее масштабирование до более крупных и сложных сред MIoT может вызвать дополнительные проблемы. Обеспечение стабильной производительности и надёжности в более крупной сети устройств может быть затруднено

III. MediHunt в сравнении с другими решениями

MediHunt выделяется среди фреймворков сетевой криминалистики, особенно в контексте медицинских сред Интернета вещей (MIoT), благодаря своей специализированной направленности, производительности и точности. При сравнении MediHunt с другими сетевыми криминалистическими фреймворками подчёркивается его уникальность и эффективность:

- **Специализированный фокус на MIoT**: В отличие от многих общих фреймворков сетевой криминалистики, MediHunt разработан специально для домена MIoT. Такая специализация позволяет ИТ-отделу решать уникальные задачи и требования, предъявляемые к медицинским устройствам Интернета вещей, таким как ограниченность ресурсов и необходимость обнаружения атак в режиме реального времени.
- **Обнаружение атак в режиме реального времени**: способность MediHunt обнаруживать атаки в режиме реального времени является значительным преимуществом. Эта функция имеет решающее значение для сред MIoT, где своевременное обнаружение может предотвратить потенциальный вред пациентам и медицинским операциям. Хотя обнаружение в режиме реального времени является целью многих фреймворков, MediHunt's адаптирована к ограниченному характеру устройств MIoT, обеспечивая минимальное влияние на производительность устройства.
- **Производительность и точность**: MediHunt демонстрирует исключительную производительность и точность при обнаружении сетевых атак. Благодаря баллам F1 и точности обнаружения, превышающей 0,99, он превосходит многие существующие платформы по своей способности точно выявлять вредоносные действия без высокого уровня ложных срабатываний. Такой уровень точности особенно важен в медицинских учреждениях, где ложные срабатывания могут иметь серьёзные последствия.
- **Эффективность использования ресурсов**: несмотря на свои широкие возможности, MediHunt разработан с учётом экономии ресурсов, что делает его подходящим для развёртывания на устройствах MIoT с ограниченными ресурсами. Это контрастирует с некоторыми другими фреймворками, которые могут требовать более значительных вычислительных ресурсов, что делает их менее жизнеспособными для развёртывания в сценариях MIoT.
- **Интеграция с машинным обучением**: MediHunt использует модели машинного обучения для расширения возможностей обнаружения атак. В то время как другие фреймворки также используют машинное обучение, подход MediHunt специально разработан для типов атак, распространенных в сетях MIoT, с использованием пользовательского набора данных, который включает данные потока как для атак уровня TCP / IP, так и для атак прикладного уровня.
- **Набор данных и обучение модели**: Пользовательский набор данных для обучения моделей машинного обучения - ещё один аспект, в котором выделяется MediHunt. Многие фреймворки сталкиваются с нехваткой всеобъемлющих наборов данных для обучения, особенно в контексте MIoT. MediHunt устраняет этот пробел, используя набор данных, охватывающий широкий спектр сценариев атак, имеющих отношение к средам MIoT

IV. Предыдущие исследования

A. *Обзор существующих систем криминалистики*

Освещаются сильные стороны и ограничения существующих фреймворков. Например, традиционные системы цифровой криминалистики хорошо зарекомендовали себя и широко использовались в различных контекстах, но они часто оказываются несостоятельными при применении к уникальным и сложным средам систем интернета вещей. Обсуждаемые фреймворки включают те, которые ориентированы на криминалистику устройств, сетевую криминалистику экспертизу и облачную криминалистику, каждая из которых имеет свой собственный набор методологий и инструментов, предназначенных для решения конкретных задач криминалистики.

B. *Проблемы криминалистики MIoT*

Подчёркиваются уникальные проблемы, с которыми сталкивается криминалистика в области медицинского интернета вещей (MIoT). Одной из основных проблем является ограниченность ресурсов устройств MIoT, которые часто имеют ограниченную вычислительную мощность, память и возможности хранения данных. Это затрудняет внедрение традиционных инструментов и методов криминалистики. Кроме того, существует значительная нехватка полных наборов данных для обучения моделей машинного обучения, которые имеют решающее значение для эффективного обнаружения атак и криминалистического анализа. Неоднородность устройств MIoT с их различными операционными системами, протоколами связи и форматами данных усложняет процесс криминалистики.

C. *Сравнение с традиционной криминалистикой*

Проводится сравнение между традиционной цифровой криминалистикой и криминалистикой Интернета вещей. Традиционная цифровая криминалистика обычно имеет дело с чётко определёнными и однородными средами, такими как персональные компьютеры и серверы, где могут быть эффективно применены стандартные инструменты и методы. Напротив, криминалистике Интернета вещей приходится иметь дело с крайне неоднородной средой и ограниченными ресурсами. Обычные инструменты криминалистики часто неадекватны для систем Интернета вещей, которые требуют специализированных подходов для работы с разнообразным и динамичным характером устройств и сетей Интернета вещей.

D. *Использование машинного обучения*

Обсуждается применение методов машинного обучения (ML) в сетевой криминалистике, в частности, для обнаружения и анализа аномалий сетевого трафика. Машинное обучение обладает значительным потенциалом для повышения точности и эффективности forensics-исследований за счёт выявления закономерностей и аномалий в сетевом трафике, которые могут указывать на вредоносную активность. Эффективность моделей ML в зависит от доступности высококачественных наборов данных, охватывающих широкий спектр сценариев атак, особенно адаптированных к характеристикам систем Интернета вещей на основе MQTT.

E. *Существующие наборы данных*

Представлен обзор существующих наборов данных, используемых для обучения моделей машинного обучения в сетевой криминалистике. Эти наборы данных имеют решающее значение для разработки и валидации ML-моделей, но они часто имеют ограничения с точки зрения разнообразия и всесторонности. Многие существующие наборы данных неадекватно отражают разнообразие сценариев атак в системах Интернета вещей на основе MQTT, что ограничивает эффективность обученных моделей. В этом разделе подчёркивается важность разработки более полных и репрезентативных наборов данных для повышения эффективности криминалистических инструментов, основанных на ML.

F. *Степень разработанности темы*

Выявляются пробелы в текущей литературе (степень научно-практической разработанности темы) по криминалистике MIoT. Одним из ключевых пробелов является потребность в возможностях обнаружения атак в режиме реального времени, которые необходимы для оперативного выявления и смягчения угроз в средах MioT и усовершенствованных методах сохранения forensics-доказательств, гарантирующих, что они останутся нетронутыми и допустимыми в ходе forensics-исследования. Устранение этих пробелов имеет решающее значение для развития области цифровой MIoT-криминалистики и повышения безопасности и надёжности медицинских систем Интернета вещей.

V. Предлагаемая система сетевой криминалистики

- **Разработка фреймворка**: MediHunt разработан для решения конкретных задач сетевой криминалистики в средах MIoT с особым упором на протокол MQTT. Он направлен на обнаружение атак в режиме реального времени и сохранение необходимых журналов для последующего анализа.
- **Обнаружение атак в режиме реального времени**: Способность обнаруживать кибератаки по мере их возникновения имеет решающее значение для уменьшения потенциального ущерба и немедленного начала forensics-исследования.
- **Механизм хранения журналов**: Учитывая ограниченность памяти устройств MIoT, MediHunt включает эффективный механизм хранения журналов, что гарантирует доступность журналов, относящихся к обнаруженным атакам, для анализа без перегрузки ёмкости хранилища.
- **Интеграция с машинным обучением**: MediHunt использует методы ML для расширения возможностей обнаружения атак. Он использует пользовательский набор данных, который включает данные потока как для атак уровня TCP / IP, так и для атак прикладного уровня, устраняя нехватку наборов данных для систем интернета вещей на основе MQTT.
- **Набор данных и обучение модели**: Пользовательский набор данных, используемый в MediHunt, охватывает широкий спектр сценариев атак, позволяя обучать модели ML распознавать различные типы кибератак. Шесть различных моделей ML были обучены и оценены на предмет их эффективности при обнаружении атак в режиме реального времени.
- **Показатели производительности**: Эффективность MediHunt количественно измеряется с использованием баллов F1 и точности обнаружения, и достигнутая высокая производительность превышает 0,99, что указывает на её надёжность при обнаружении атак в сетях MQTT.
- **Комплексный криминалистический анализ**: помимо обнаружения атак, MediHunt облегчает процесс комплексного анализа. Он поддерживает сбор, анализ, представление и сохранение цифровых доказательств в соответствии с принципами сетевой криминалистики.
- **Эффективность использования ресурсов**: MediHunt разработан с учётом экономии ресурсов, что делает его подходящим для развёртывания на устройствах MIoT с ограниченными ресурсами.

VI. Обучение модели ML

A. *Сбор данных о сетевом трафике MQTT*

- **Типы собираемых данных**: Собираемые данные включают как обычный трафик, так и трафик атаки. Это гарантирует, что набор данных является

всеобъемлющим и может быть использован для эффективного обучения моделей машинного обучения.

- **Данные на основе потоков**: сбор данных на основе потоков включает информацию о потоках связи между устройствами. Этот тип данных имеет решающее значение для обнаружения аномалий и атак в сетевом трафике.
- **Сценарии атак**: сценарии произвольных атак моделируются для генерации атакующего трафика и включают атаки TCP / IP и прикладного уровня, специфичные для MQTT.
- **Генерация набора данных**: Собранные данные обрабатываются для создания набора данных, который может быть использован для обучения моделей машинного обучения. Этот набор данных включает помеченные экземпляры как обычного трафика, так и трафика атаки.

B. Обучение модели ML и анализ эффективности

- **Модели машинного обучения**: оцениваются шесть различных моделей, включая деревья принятия решений, случайные леса, машины опорных векторов и нейронные сети.
- **Процесс обучения**: Процесс обучения включает использование сгенерированного набора данных для обучения моделей машинного обучения. Модели обучены распознавать закономерности в данных, которые указывают на нормальный трафик или трафик атаки.
- **Показатели производительности**: Производительность обученных моделей оценивается с использованием таких показателей, как оценка F1 и точность обнаружения, которые обеспечивают количественный показатель эффективности моделей при обнаружении атак.
- **Высокая производительность**: достигаются баллы F1, а точность обнаружения превышает 0.99, что подтверждает эффективность обнаружения атак в режиме реального времени.
- **Обнаружение в режиме реального времени**: t обученные модели интегрированы в платформу MediHunt для обеспечения обнаружения атак в режиме реального времени. Это позволяет незамедлительно реагировать и смягчать потенциальные угрозы.

VII. Оценка на Raspberry Pi

- **Реализация на Raspberry Pi**: проанализирована производительность алгоритмов машинного обучения (ML) на моделях Raspberry Pi 3B для реализации платформы сетевой криминалистики MediHunt на устройствах MIoT с ресурсами.
- **Сопоставимое время вывода и обучения**: Оценка показала, что время вывода и обучения алгоритмов ML были сопоставимы на устройствах Raspberry Pi. В частности, время вывода на облачной платформе составляло около 2 мс, в то время как на Raspberry Pi оно составляло 0,17 мс.
- **Легковесная обнаружения вторжений**: MediHunt описывается как облегчённое решение для обнаружения вторжений, разворачиваемое на ограниченных ресурсами устройствах (Raspberry Pi).
- **Обнаружение атак в режиме реального времени**: подчёркивается способность платформы обнаруживать атаки в режиме реального времени, обеспечивая немедленное реагирование и смягчение потенциальных угроз.
- **Эффективное использование ресурсов**: несмотря на широкие возможности для сетевой криминалистики, платформа MediHunt разработана с учётом экономии ресурсов, что делает её подходящей для развёртывания на устройствах MIoT с ограниченными ресурсами, таких как Raspberry Pi.