Abstract – The paper "MediHunt: A Network Forensics Framework for Medical IoT Devices" presents the development of MediHunt framework designed for real-time detection of network flow-based traffic attacks in MQTT networks, which are commonly used in smart hospital environments. MediHunt can detect a variety of TCP/IP layers and application layer attacks on MQTT networks by leveraging machine learning models. The framework aims to enhance the forensic analysis capabilities in MIoT environments, ensuring effective tracing and mitigation of malicious activities.

I. Introduction

The paper "MediHunt: A Network Forensics Framework for Medical IoT Devices" addresses the need for robust network forensics in Medical Internet of Things (MIoT) environments, particularly focusing on MQTT (Message Queuing Telemetry Transport) networks. These networks are commonly used in smart hospital environments for their lightweight communication protocol. It highlights the challenges in securing MIoT devices, which are often resource-constrained and have limited computational power. The lack of publicly available flow-based MQTT-specific datasets for training attack detection systems is mentioned as a significant challenge.

The paper presents MediHunt as an automatic network forensics solution designed for real-time detection of network flow-based traffic attacks in MQTT networks. It aims to provide a comprehensive solution for data collection, analysis, attack detection, presentation, and preservation of evidence. It is designed to detect a variety of TCP/IP layers and application layer attacks on MQTT networks. It leverages machine learning models to enhance the detection capabilities and is suitable for deployment on resource constrained MIoT devices.

The primary objective of the MediHunt is to strengthen the forensic analysis capabilities in MIoT environments, ensuring that malicious activities can be traced and mitigated effectively.

II. Benefits and Drawbacks of Proposed Solution

A. *Benefits*

- **Real-time Attack Detection**: MediHunt is designed to detect network flow-based traffic attacks in real-time, which is crucial for mitigating potential damage and ensuring the security of MIoT environments.
- **Comprehensive Forensic Capabilities**: The framework provides a complete solution for data collection, analysis, attack detection, presentation, and preservation of evidence. This makes it a robust tool for network forensics in MIoT environments.
- **Machine Learning Integration**: By leveraging machine learning models, MediHunt enhances its detection capabilities. The use of a custom dataset that includes flow data for both TCP/IP layer and application layer attacks allows for more accurate and effective detection of a wide range of cyber-attacks.
- **High Performance**: The framework has demonstrated high performance, with F1 scores and detection accuracy exceeding 0.99 and indicates that it is highly reliable in detecting attacks on MQTT networks.
- **Resource Efficiency**: Despite its comprehensive capabilities, MediHunt is designed to be resource-efficient, making it suitable for deployment on resource-constrained MIoT devices like Raspberry Pi.

B. *Drawbacks*

- **Dataset Limitations**: While MediHunt uses a custom dataset for training its machine learning models, the creation and maintenance of such datasets can be challenging. The dataset needs to be regularly updated to cover new and emerging attack scenarios.
- **Resource Constraints**: Although MediHunt is designed to be resource-efficient, the inherent limitations of MIoT devices, such as limited computational power and memory, can still pose challenges. Ensuring that the framework runs smoothly on these devices without impacting their primary functions can be difficult.
- **Complexity of Implementation**: Implementing and maintaining a machine learning-based network forensics framework can be complex. It requires expertise in cybersecurity and machine learning, which may not be readily available in all healthcare settings.
- **Dependence on Machine Learning Models**: The effectiveness of MediHunt heavily relies on the accuracy and robustness of its machine learning models. These models need to be trained on high-quality data and regularly updated to remain effective against new types of attacks.
- **Scalability Issues**: While the framework is suitable for small-scale deployments on devices like Raspberry Pi, scaling it up to larger, more complex MIoT environments may present additional challenges. Ensuring consistent performance and reliability across a larger network of devices can be difficult

III. MediHunt vs other frameworks

MediHunt stands out among network forensics frameworks, particularly in the context of Medical Internet of Things (MIoT) environments, due to its specialized focus, performance, and accuracy. When comparing MediHunt to other network forensics frameworks, several key aspects highlight its distinctiveness and effectiveness:

- **Specialized Focus on MIoT**: Unlike many network forensics frameworks, MediHunt is specifically designed for the MIoT domain. This specialization allows it to address the unique challenges and requirements of medical IoT devices, such as resource constraints and the need for real-time attack detection.
- **Real-time Attack Detection**: MediHunt's capability to detect attacks in real-time is a significant advantage. This feature is crucial for MIoT environments where timely detection can prevent potential harm to patients and healthcare operations. MediHunt's implementation is tailored to the lightweight nature of MIoT devices, ensuring minimal impact on device performance.
- **Performance and Accuracy**: MediHunt demonstrates exceptional performance and accuracy in detecting network attacks. With F1 scores and detection accuracy exceeding 0.99, it surpasses many existing frameworks in its ability to accurately identify malicious activities without a high rate of false positives. This level of accuracy is particularly important in healthcare settings, where false alarms can have serious implications.
- **Resource Efficiency**: Despite its comprehensive capabilities, MediHunt is designed to be resource-efficient, making it suitable for deployment on resource-constrained MIoT devices. This contrasts with some other frameworks that may require more substantial computational resources, making them less viable for deployment in MIoT scenarios.
- **Machine Learning Integration**: MediHunt leverages machine learning models to enhance its attack detection capabilities. While other frameworks also use machine learning, MediHunt's approach is specifically tuned for the types of attacks prevalent in MIoT networks, using a custom dataset that includes flow data for both TCP/IP layer and application layer attacks.
- **Dataset and Model Training**: The custom dataset for training machine learning models is another aspect where MediHunt stands out. Many frameworks struggle with the lack of comprehensive datasets for training, especially in the context of MIoT. MediHunt addresses this gap by leveraging a dataset that covers a wide range of attack scenarios relevant to MIoT environments

IV. Related Work

A. Overview of Existing Forensic Frameworks

. This review highlights the strengths and limitations of existing network forensic frameworks and their applications across different domains. For instance, traditional digital forensics frameworks are well-established and have been extensively used in various contexts, but they often fall short when applied to the unique and complex environments of IoT systems. The frameworks discussed include those that focus on device forensics, network forensics, and cloud forensics, each with its own set of methodologies and tools designed to address specific forensic challenges.

B. Challenges in MIoT Forensics

The section emphasizes the unique challenges faced in Medical Internet of Things (MIoT) forensics. One of the primary challenges is the resource constraints of MIoT devices, which often have limited computational power, memory, and storage capabilities. This makes it difficult to implement traditional forensic tools and techniques. Additionally, there is a significant lack of comprehensive datasets for training machine learning models, which are crucial for effective attack detection and forensic analysis. The heterogeneity of MIoT devices, with their varied operating systems, communication protocols, and data formats, complicates the forensic process.

C. Comparison with Traditional Forensics

A comparison is made between traditional digital forensics and IoT forensics. Traditional digital forensics typically deals with well-defined and homogeneous environments, such as personal computers and servers, where standard tools and techniques can be effectively applied. In contrast, IoT forensics must contend with a highly heterogeneous and resource-constrained environment. Conventional forensic tools are often inadequate for IoT systems, which require specialized approaches to handle the diverse and dynamic nature of IoT devices and networks.

D. Use of Machine Learning

The section discusses the application of machine learning (ML) techniques in network forensics, particularly for detecting and analyzing network traffic anomalies. Machine learning offers significant potential for improving the accuracy and efficiency of forensic investigations by identifying patterns and anomalies in network traffic that may indicate malicious activity. However, the effectiveness of ML models depends heavily on the availability of high-quality datasets that cover a wide range of attack scenarios. The need for specific datasets tailored to the characteristics of MQTT-based IoT systems is particularly highlighted.

E. Existing Datasets

A review of existing datasets used for training machine learning models in network forensics is provided. These datasets are critical for developing and validating ML models, but they often have limitations in terms of diversity and comprehensiveness. Many existing datasets do not adequately represent the variety of attack scenarios that can occur in MQTT-based IoT systems, which limits the effectiveness of the trained models. The section underscores the importance of developing more comprehensive and representative datasets to improve the performance of ML-based forensic tools.

F. Gap in Literature

Finally, the section identifies gaps in the current literature on MIoT forensics. One of the key gaps is the need for real-time attack detection capabilities, which are essential for promptly identifying and mitigating threats in MIoT environments. Additionally, there is a need for improved methods for

preserving forensic evidence, ensuring that it remains intact and admissible in legal proceedings. Addressing these gaps is crucial for advancing the field of MIoT forensics and enhancing the security and reliability of medical IoT systems.

V. Proposed Network Forensics Framework

- **Framework Design**: MediHunt is designed to address the specific challenges of network forensics in MIoT environments, particularly focusing on the MQTT protocol. It aims to detect attacks in real-time and preserve the necessary logs for forensic analysis.
- **Real-time Attack Detection**: Capability to detect cyber-attacks as they happen is crucial for mitigating potential damage and for the immediate initiation of forensic analysis.
- **Log Storage Mechanism**: Given the memory constraints of MIoT devices, MediHunt incorporates an efficient log storage mechanism. It ensures that logs relevant to detected attacks are stored for further analysis without overwhelming the storage capacity.
- **Machine Learning Integration**: MediHunt leverages ML techniques to enhance its attack detection capabilities. It utilizes a custom dataset that includes flow data for both TCP/IP layer and application layer attacks, addressing the lack of datasets for MQTT-based IoT systems.
- **Dataset and Model Training**: The custom dataset used in MediHunt covers a wide range of attack scenarios, enabling the training of ML models to recognize various types of cyber-attacks. Six different ML models were trained and evaluated for their effectiveness in real-time attack detection.
- **Performance Metrics**: MediHunt's effectiveness is quantitatively measured using F1 scores and detection accuracy and achieved high performance exceeding 0.99, indicating its reliability in detecting attacks on MQTT networks.
- **Comprehensive Forensic Analysis**: Beyond attack detection, MediHunt facilitates a comprehensive forensic analysis process. It supports the collection, analysis, presentation, and preservation of digital evidence, adhering to principles of network forensics.
- **Resource Efficiency**: MediHunt is designed to be resource-efficient, making it suitable for deployment on resource-constrained MIoT devices.

VI. ML Model Training

A. *MQTT Network Traffic Data Collection*

- **Types of Data Collected**: The data collected includes both normal and attack traffic. This ensures that the dataset is comprehensive and can be used to train machine learning models effectively.
- **Flow-Based Data**: collecting flow-based data includes information about the communication flows between devices. This type of data is crucial for detecting anomalies and attacks in network traffic.
- **Attack Scenarios**: various attack scenarios are simulated to generate attack traffic and include TCP/IP and application layer attacks specific to the MQTT.
- **Dataset Generation**: The collected data is processed to generate a dataset that can be used for training machine learning models. This dataset includes labeled instances of both normal and attack traffic.

B. *ML Model Training and Performance Analysis*

- **Machine Learning Models**: Six different models are evaluated, including decision trees, random forests, support vector machines, and neural networks.
- **Training Process**: The training process involves using the generated dataset to train the machine learning models. The models are trained to recognize patterns in the data that indicate normal or attack traffic.
- **Performance Metrics**: The performance of the trained models is evaluated using metrics such as F1 score and detection accuracy that provide a quantitative measure of the models' effectiveness in detecting attacks.
- **High Performance**: achieved with F1 scores and detection accuracy exceeds 0.99 that indicates the highly effectiveness in detecting attacks in real-time.
- **Real-Time Detection**: the trained models are integrated into the MediHunt framework to enable real-time detection of attacks. This allows for immediate response and mitigation of potential threats.

VII. Evaluation on Raspberry Pi

- **Implementation on Raspberry Pi**: The authors analyzed the performance of machine learning (ML) algorithms on Raspberry Pi 3B models to implement the MediHunt network forensics framework on resource limited MIoT devices.
- **Comparable Inference and Training Times**: The evaluation revealed that the inference and training times of the ML algorithms were comparable on the Raspberry Pi devices. Specifically, the inference time on the cloud platform was around 2ms, while on the Raspberry Pi, it was 0.17ms.
- **Lightweight Intrusion Detection System**: MediHunt is described as a lightweight intrusion detection system solution that can be readily deployed on resource constrained MIoT devices like Raspberry Pis.
- **Real-time Attack Detection**: The framework's ability to detect attacks in real-time is highlighted, enabling immediate response and mitigation of potential threats.
- **Efficient Resource Utilization**: Despite its comprehensive capabilities for network forensics, the MediHunt framework is designed to be resource-efficient, making it suitable for deployment on resource-constrained MIoT devices like Raspberry Pis.