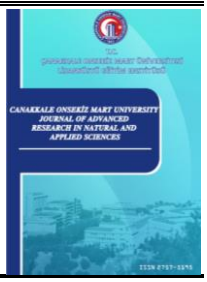
Çanakkale Onsekiz Mart University

Journal of Advanced Research in Natural and Applied Sciences

dergipark.org.tr/tr/pub/jarnas

Open Access



e-ISSN: 2757-5195 2024, Vol. 10, Issue 4, Pages: 899-912 doi.org/10.28979/jarnas.1559652

A Novel Approach for Detection of Cyber Attacks in MQTT-Based IIoT Systems Using Machine Learning Techniques

Serkan Gönen[1]

[1]Department of Software Engineering, Faculty of Engineering and Architecture, İstanbul Gelişim University, İstanbul, Türkiye

Article Info

Received: 01 Oct 2024

Accepted: 28 Nov 2024

Published: 31 Dec 2024

Research Article

Abstract − The Internet of Things (IoT) and the Industrial Internet of Things (IIoT) have grown significantly in the last decade, underlining the increasing need for effective, secure, and reliable data communication protocols. The widely accepted Message Queuing Telemetry Transport (MQTT) protocol, with its structure that meets the needs of welding-oriented devices in IoT and IIoT applications, is a prime example. However, its user-friendly simplicity also makes it susceptible to threats such as Dispersed Services Rejection (DDOS), Brete-Force, and incorrectly shaped package attacks. This article introduces a robust and reliable framework for preventing and defending against such attacks in MQTT-based IIoT systems based on the theory of merging attacks. The expert system incorporates the Adaboost model and can detect anomalies by processing network traffic in a closed setting and identifying impending threats. With its robust design, the system was subjected to various attack scenarios during testing, and it consistently detected interventions with an average accuracy of 92.7%, demonstrating its potential for use in intervention detection systems. The research findings not only contribute to the theoretical and practical concerns about the effective protection of IIoT systems but also offer hope for the future of cybersecurity in these systems.

Keywords − *IoT, IIoT, MQTT, cyber security, machine learning*

1. Introduction

In the past few years, systems of the Internet of Things (IoT) and the so-called Industrial Internet of Things (IIoT) have rapidly developed technological paradigms. On the one hand, IoT is more like a network structure where all physical devices are connected to the internet and can share data. At the same time, IIoT firmly takes that technology one step further in automating and enhancing industrial activities in smaller terraces. Fast and better protocols for data exchange between the devices are indispensable if one is to succeed in such systems. In this regard, the telemetry transmission protocol, the MQQT, has replaced the most popular data exchange protocol in IoT and IIoT systems. The MQTT protocol is a variation that provides an interesting solution mainly for low-end devices due to its low resource profiles and bandwidth consumption. This is important, especially in terms of energy efficiency and transmission of data over a network with minimum overhead. However, many cyber security risks exist because of the protocol's simplicity and lightweight. Particularly, crimes of this sort include brute-force attacks, Distributed Denial of Service (DDoS) attacks whereby the target system is flooded with incoming packets from uncontrolled users and malformed packets that attack the application layer of the systems using the MQTT protocol. In this regard, several security measures and intrusion detection systems are being developed to better secure systems based on the MQTT protocol.

[1]sgonen@gelisim.edu.tr (Corresponding Author)

Numerous works of literature aim to prevent the ambiguity of the security issues related to the MQTT protocol. Al-Fayoumi and Al-Haija [1] achieved approximately a 99.5% success rate in detecting low-rate DDoS attacks in MQTT-based IoT environments using various machine learning models, primarily decision trees. Due to the results presented in this paper, it is evident how helpful machine learning algorithms are during attempts to perform an active defense. Simultaneously, Fikriansyah et al. achieved a 95% accuracy rate against DDoS attacks in IIoT environments, which leveraged the Random Forest algorithm to identify the attacks. Considering these findings, it can be reasoned that such attacks against the MQTT protocol employing machine learning devices are easily identifiable [2]. Nevertheless, most of this study has mainly focused on classes of attacks, and it is necessary to broaden the scope of studies. There is a greater demand for real-time detection systems, especially in IIoT. Some IIoT threats involve the latency of threats and detection effectiveness. In inter-industrial IoT connections, potential threats must always be detected and attended to with speed and precision since either dawning or loss of data could halt production or any other proceedings.

This research seeks to create an expert system based on MQTT capable of identifying and guarding against cyber-attacks in IIoT systems. This expert system processes the network traffic in near real-time using machine learning techniques and identifies possible risks. The primary goal of the research is to find a security solution that maintains the performance of IIoT systems while raising their security level. Within this framework, experiments were performed on different kinds of attacks, including brute-force, DDoS, and malicious packet attacks, and machine-learning algorithms were built on the obtained data. In this phase, attacks were performed to gather data to be used in the expert system, and such data was used to monitor network traffic. This system could perform IDS attacks after the learning processes on this data had been completed and all relevant results were analyzed. This study provides many types of attack examples that can be performed, and it provides real-time interaction and reaction. Moreover, in light of the area for improvement, the expert system is expected to be able to meet threats within its scope later on. In such a case, the present study's findings benefit academic and industrial uses and implications.

The present study concerns the cyber security risks in such an MQTT-based IIoT system, draws attention to issues of this kind, and offers new solutions. More precisely, the expert system constructed using the AdaBoost algorithm has high accuracy and less resource utilization, which is an essential advancement in developing real-time detection and preventing intrusions in IIoT systems. This research serves as a critical illustration of the application of technologies such as machine learning and expert systems and, at the same time, offers practical approaches to the problem of security in an industrial context. For this reason, it impacts both the academic and industrial settings and is worth viewing as a valuable addition to the existing body of knowledge on securing IIoT systems.

Incorporating cryptographic models in machine learning has emerged as a pivotal strategy for enhancing data security in IoT and IIoT systems. Techniques such as homomorphic encryption allow machine learning algorithms to operate on encrypted data without decryption, safeguarding data confidentiality throughout training and inference. Integrating these cryptographic models can fortify IIoT systems against data exfiltration and unauthorized access while maintaining operational efficiency.

Within the article, we present how security silos of the systems integrating MQTT protocol are built starting MMA, from generalization and application of Security areas to a review of literature on existing Security solutions targeted at this emerging Security area. Then, it concentrates on the expert system design process with the help of machine learning, the performance analysis, the description of the benefits and the limitations of the system, and the suggestion for further work.

2. Literature Research

Due to its lightweight, practical nature, the MQTT protocol is gaining popularity in the transmission of data on the Internet of Things (IoT) and Industrial Internet of Things (IIoT) systems. Nevertheless, this protocol has been a target for cyber security research because of its risks. Al-Fayoumi and Al-Haija [1] utilized decision

trees and machine learning models. They achieved high outcomes in detecting low-rate distributed denial of service attacks (LR-DDoS) within MQTT IoT environments. Likewise, Fikriansyah et al. [2] used a Random Forest algorithm to detect DDoS attacks and reported encouraging results in IIoT environments. At the same time, Shahri et al. [3] proposed and implemented an SDN-based MQTT framework for enhanced real-time performance and scalability within industrial scenarios. Mahajan et al. [4] elevated communication security through symmetric encryption algorithms such as AES and HMAC over the MQTT protocol, from ensuring data integrity to data confidentiality. Moreover, Kombate et al. [5] addressed the weaknesses in the MQTT protocol and provided cyber-range technologies for mitigating cyber threats, including replay attacks, MITM, and DDoS. Buccafurri et al. [6] also presented the OTP verification over blockchain technology to reduce energy consumption and improve security during MQTT communication. It should be noted that, as much as the MQTT protocol facilitates data transfer in IoT and IIoT systems, such systems may be rendered susceptible because of security loopholes.

More specifically, applying algorithms like the AES and HMAC solves problems like data eavesdropping and data tampering, as this guarantees the confidentiality and integrity of the data in question. This study shifts the existing paradigm of MQTT, which relies on security over the resources it hosts, and aims to maintain secure MQTT communication with optimal security resources on devices. Liu et al. [7] further developed that and suggested the adoption of more SM2 and SM4 Cryptographic algorithms along with the MQTT protocol, extending the functionalities of user validation and access control. This scheme fills the conventional emptiness in the protocol's security vulnerabilities, mainly in data security and mutual authentication. These results emphasize that these algorithms do improve security and do not interfere with the efficiency of the systems to a significant extent. Saqib and Moon [8] have also provided lightweight solutions in the form of multi-factor authentication systems for edge computing. A new security architecture was presented for IoT applications based on the MQTT protocol, which allowed the authors to use ECC and fuzzy extractors. This system protects IoT devices from cryptographic attacks by enhancing session essential consent and reciprocated security arrangements. It is more suitable for low-powered people who work with less computation.

As another approach, Koprov et al. [9] proposed a framework for machine authentication for the MQTT 5.0 standard, with the support of functional data analysis (FDA), among others. Through deeper machine data analysis, this technique aims to raise the bar for better machine authentication within IIOT and Smart Manufacturing. This methodology raises confidence in the authentication processes, especially in the digital twins and predictive maintenance domains. Lohachab and Karambir [10] presented one lightweight authentication and authorization method for IoT device Communications using ECC, intended for resource-limited regional IoT devices. This development helps improve IoT device authentication mechanisms by allowing cross-device safe content sharing. In particular, the study attempts to provide such optimal security deployments and focuses on deployment challenges posed by resource-constrained devices. IoT and IIOT systems have created various frameworks and methodologies to solve safety problems caused by these interdependent devices, mainly as they are used more widely in critical infrastructures. Vaccari et al. [11] in MQTTSET suggested that the data set be designed to form a basis for machine learning models to protect MQTT-based IoT networks from possible cyber-attacks. In light of this, Mishra and Kertesz [12] have reviewed the existing literature on using MQTT in IoT systems. Investigation emphasized the aspects in which rapid growth areas and safety measures grow as fast as protocol use. Francis et al. [13] have a genetic algorithm and random forest-based attack detection system (GA_RF), which obtains almost perfect accuracy in detecting threats in MQTT-enabled IIOT systems. To expand their research, Patel and Doshi [14] developed a new MQTT protocol safety standard that increases user authentication and data safety on IoT networks. Deng [15] proposes spectrogram analysis and an attack detection system that integrates spectrogram analysis and evolving neural networks to better detect interventions in the context of IoT. Prajisha and Vasudevan [16] proposed a better performance of the Intruder detection system using a slight Gradian reinforcement machine through chaos salp herd optimization. Likewise, investigations were conducted on which defense measures could be applied to protect IIOT devices from attacks.

According to Karacaymaz and Artuner [17], the expert system based on AI should be unique. Selamnia et al. [18] aimed to develop a comprehensive cybersecurity training program, including course content development, recruitment, and system setup, utilizing instructors and mentors to deliver practical and theoretical cybersecurity skills over ten months. Zuhari et al. [19] MQTT networks were improved in property engineering for a real-time pseudonym for a relief-based routing protocol. They presented the usefulness of some tree models in perception performance. Sharma and Bhushan propose a hybrid framework incorporating PUFs and machine learning to defend against distributed denial of service (DDoS) attacks in MQTT-based IoT systems. The system improves upon existing credential-based authentication by augmenting the PUF-enabled mechanism to create real-time, in-depth optimization. It combines it with a machine learning-based Intrusion Detection System (IDS), which aims to eliminate malevolent attacks against the system assets at the time of their occurrence. Their findings are not only secure but also efficient in terms of resource utilization, which makes them a better vendor for usage on lightweight IoT devices [20].

This study provides valuable insights into new tendencies toward creating security systems, especially those dedicated to addressing the issue of energy use in IoT devices. In general, the gaps covered by collaboration and some available studies included approaches that raised the level of security on the IIoT systems built on MQTT, including machine learning and encryption technologies. However, to address the issues of real-time detection and flexibility concerning the different aspects of the attacks, the following section describes the expert system developed to enhance security within these environments.

3. Methodology

This work focused on creating and applying an expert system based on the MQTT protocol for determining and preventing cyber threats to IIoT systems. This methodology combined classical network defense approaches with contemporary machine learning approaches to solve the security problem of IIoT architectures.

3.1. Overview of the Methodology

The methodology can be broken down into several phases illustrated in Figure 1. This method can be further shown in its three main stages: selecting a machine learning model, including attack simulation and data collection, expert system design, and evolution of the expert system. The first stage included a simulated attack of various attacks against the previously described test bed. At this stage, brute force, DDOs, and poorly formatted package attacks were carried out on the MQTT broker. During these attacks, data on network traffic were then caught to provide a database useful for training machine learning models. Roller was appointed and synthesized by expert systems. This was the background of the methodology for creating an expert system to analyze real-time network traffic and detect potential attacks. Therefore, this expert system was created using machine learning methods to classify network traffic into normal and attack. The system was well planned, aiming at high efficiency and accuracy in detecting attacks to achieve a minimum delay in the network.

The machine learning model selection process was integral to the expert system's critical development. The expert system can try simple machine-learning approaches: decision trees, random forests, and the AdaBoost algorithm. After considering the algorithms' performance, the expert system concluded that the AdaBoost model best determines cyber-attacks in MQTT-based IIOT systems.

Training and Verification: After selecting the Adaboost model, training was done against the data set collected from the test bed. The data set was prepared at 70% and 30% for training and verification. This division helped make the model more useful for practical use by enabling the model to apply what it has learned to 'new' and 'invisible' data. The model was taught to identify the different attack classes in network traffic, for example, the traffic models of the explosive DDOs and the misshaped package corresponding to the packaged package structures.

Assessment Metrics: To determine the results of the expert system, performance, accuracy, sensitivity, remembering, and F-score, but not limited to them, were evaluated with various evaluation indices. It was possible to provide an unmistakable appearance on the healing side of the system about systematic attacks and false positives. The Adaboost model was extraordinary in successfully identifying cyber-attacks with 92.7 % accuracy in these tests.

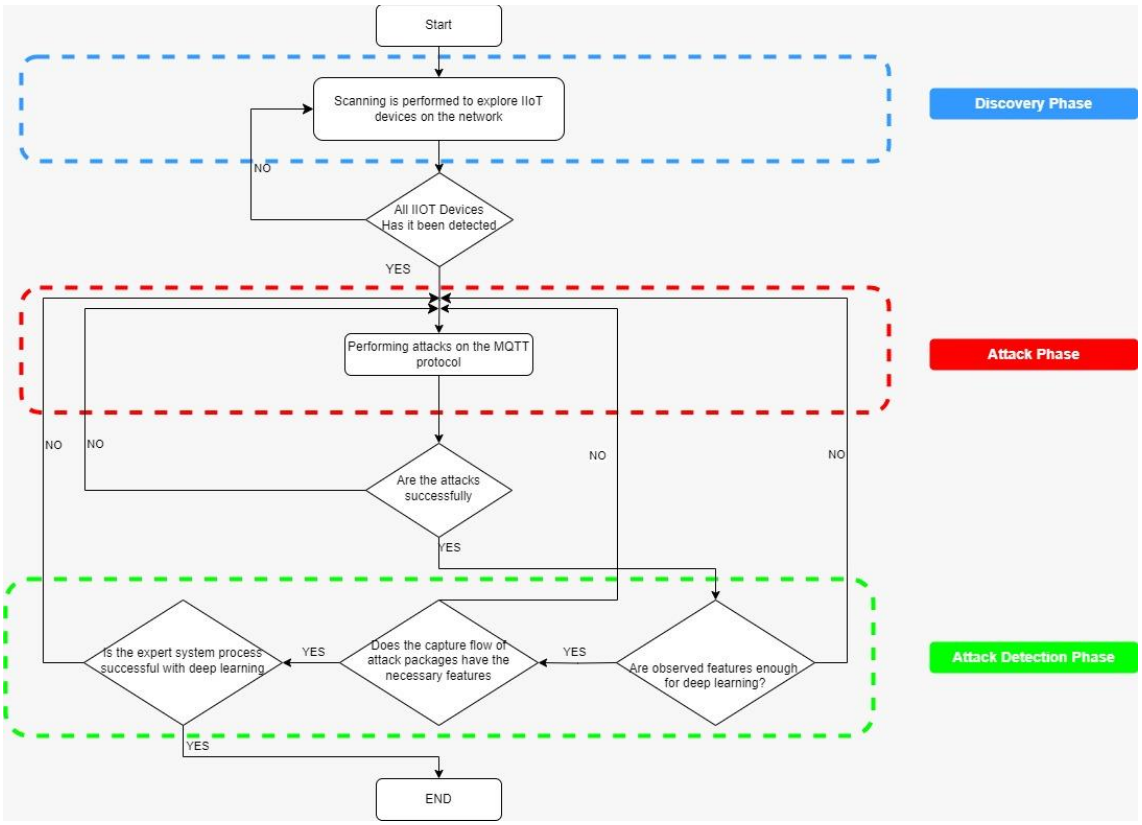


Figure 1. Workflow diagram

3.2. Algorithm Comparison and Selection

The approach also included a comprehensive comparison of some machine learning approaches. Decision trees and random forest approaches were among those who were considered promising. However, regarding capturing additional forms such as Bute Force attacks, Adaboost uses many weak classes, making them a single robust classification, thus increasing the overall determination accuracy. Expert system.

3.3. Real-time Detection and Visualization

One element of the expert system was to fulfill the analysis purpose and explain that the attack sets would be fast during attacks while answers would be obtained. The system's ability to provide real-time detection was more evaluated during test bed assessments. Even in high-level stress tests interacting with many devices simultaneously, the expert system can detect attacks without much delay against the network. A visualization tool was created as part of the methodology to improve the analysis and understanding of system outputs. The tool reported the output of expert system analysis: what kind of attacks were defined and how the system reacted over time. Graphics were made to demonstrate the number of attacks, the conditions in which the system's perception works effectively, and how the change in the attack affects the system's perception rate.

3.4. Increased Learning and Evolution of the Expert System

Finally, the methodology has been provided with a certain flexibility so that the expert system can learn and improve its performance over time. More data are obtained and processed, and machine learning models can

be better re-trained for accuracy and flexibility in new attack styles. Since it is a dynamic environment with a changing threat view, continuous learning ability is necessary for IIOT.

4. Testbed

In this study, an advanced test bed was developed instead of just simulating the environment (IIOT) to concentrate on evaluating the safety of the telemetry message transmission protocol, referred to as the MQTT protocol in this work. Another protocol often utilized in piggybacked IoT systems is the MQTT protocol. It is famous for its lightweight and efficiency, whereas other protocols may not function well, and it is increasingly used in IIOT applications. Such simplicity can also be a weakness, as cyber criminals can utilize it through brute attacks, DODS, and malformed packet attacks.

The test environment was meticulously constructed to replicate a realistic MQTT-based IIoT system, encapsulating essential components such as the MQTT broker, simulated IoT devices, and network monitoring tools. The MQTT broker acted as the central communication hub, facilitating data exchange between publishers (data-generating devices) and subscribers (data-receiving applications). The broker's susceptibility to attacks, including brute force, DDoS, and malformed packet injections, was tested to analyze the system's resilience and behavior under controlled conditions.

4.1. Components of the Testbed

The testbed was meticulously designed to incorporate all the critical elements in a typical IIoT system that uses the MQTT protocol. It simulates a networked environment where various devices communicate through an MQTT broker, which facilitates the exchange of messages between publishers (devices generating data) and subscribers (applications or devices that consume the data).

MQTT Broker: The broker is the backbone of the testbed, acting as the intermediary between IIoT devices. It ensures data integrity and delivery between the publishers and subscribers. In this setup, the broker was exposed to various types of cyberattacks, simulating a real-world scenario where attackers target a central communication hub to disrupt the operation of IIoT systems.

IoT Devices: The testbed simulated various IoT devices, representing IIoT sensors and actuators typically found in industrial applications. These devices generate data that is communicated through the broker. The devices were not compromised in this scenario, but the focus was on how the broker and the communication infrastructure handled the data under attack conditions.

Attack Simulation: Various tools were used to simulate attacks on the testbed. These included brute-force tools for testing the broker's authentication robustness and DDoS tools that generated flood attacks aimed at overwhelming the broker's capacity. In addition, malformed packets were injected into the communication stream to evaluate how the broker handles corrupted data.

Network Traffic Monitoring and Data Collection: The network traffic generated during these attacks was collected and mirrored to a separate analysis system. This data collection was critical for training machine learning models and building an expert system capable of identifying patterns in the network traffic that correspond to different types of attacks. Packet data was collected from legitimate traffic and simulated attacks, ensuring that the machine learning system had a comprehensive dataset for analysis.

4.2. Attack Scenarios

The testbed was designed to accommodate several types of attacks that are commonly encountered in MQTT-based IIoT systems:

Brute-force Attacks: These attacks target the MQTT broker's authentication system, attempting to gain unauthorized access by systematically trying different username and password combinations. The testbed used

a brute-force attack as a starting point, simulating an attacker's initial attempt to penetrate the system by accessing the broker without authorization.

DDoS Attacks: DDoS attacks were simulated by generating a flood of traffic aimed at the MQTT broker, overloading its capacity to handle requests. These attacks are particularly damaging in IIoT environments, where downtime or data loss could severely impact critical industrial operations.

Malformed Packet Attacks: In addition to flooding the broker with legitimate traffic, attackers can send malformed or corrupted packets that may cause the broker to crash or behave unpredictably. These attacks were simulated by deliberately corrupting the packet data sent to the broker, testing how well the broker can handle abnormal inputs without failing.

4.3. Testbed Performance and Scalability

The testbed was one of many areas tested for the ability to ensure scalability, which was the primary task of mirroring IIoT systems' behavior in more extensive industrial settings. Every IIoT device still connects and communicates over the network using the MQTT protocol, even as the network grows into thousands of devices. The MQTT broker must cope with this immense network traffic without compromising security. This testbed was created with a view of this reality with an increasing number of devices communicating through the broker even as various attack scenarios were being launched.

Scalability could also be evaluated by gradually increasing the volume of legitimate traffic and introducing malicious packets to see how the broker performed under varied loads. This method helped understand the extent to which the volume could be increased from the basic observation made when the system was not busy with network junk traffic attacking the broker and how effective the experts would be in system attack detection as data volume escalates.

4.4. Data Collection for Machine Learning

The data obtained from the testbed was fundamental for training machine learning models, which would later be utilized within the expert system. This dataset acquired included standard traffic patterns and traffic generated during brute force attacks, DDoS, and malformed packet attacks. The variety of the data was critical to encourage sufficient various types of machine learning algorithms to allow all typical trojan systems to operate even in heavy loads. Data collected during the test phase included a comprehensive mix of legitimate and attack-specific traffic patterns. Network packets captured during brute-force, DDoS, and malformed packet scenarios were mirrored to a dedicated analysis system for further processing. The dataset incorporated diverse traffic instances to train machine learning models effectively. The dataset was divided into two parts so that the system's accuracy could be further assured: 70% was allocated for training the machine learning models, and the other 30% was kept for testing and validation. This approach helped the system extend its learning to new data it had never encountered before, making it more efficient and dependable in practical contexts.

Despite the robust data collection approach, the dataset had limitations that should be acknowledged. The primary constraint was the variety of simulated attack types. Real-world scenarios often exhibit more complex, multi-layered threats beyond the testbed's scope. The dataset size, although sufficient for this study's objectives, may not encompass the full breadth of potential cyber-attacks encountered in expansive IIoT ecosystems. Addressing these gaps by expanding the dataset to include broader attack types and higher volumes would enhance the system's robustness and real-world applicability. In summary, the testbed ensured realism and scalability for the security assessment of the IIoT system based on MQTT. Various attacks were simulated, a reasonable amount of network traffic was collected, and machine learning models were built that were destined to become part of the expert system later on. Testbeds were important as they could create real-

world environments, which helped analyze how effective various ML algorithms would be in combating cyberattacks within IIoT system setups.

5. Understanding and Mitigating Attacks on MQTT in IIoT with Expert Systems

Analysis of Attacks on MQTT in IIoT the above section covers the detection of brute force attacks, DDoS (DoS, Flood, Slowite), and malformed packet attacks on the MQTT Protocol developed for use in IIoT Supported by a core expert system. Towards this end, a brute force attack was conducted to gain unauthorized access to the MQTT Broker, which was legal user access, followed by addressing DoS and malformed attacks on the said system. Packets for the said attacks would be acquired through mirroring techniques and fed into machine learning algorithms.

Together with the section above, the results pages of the data mining software concerning the analysis are displayed in the second section. Figure 2 illustrates that each time an attack log compiled for the MQTT protocol analysis under the scope of attack analysis is uploaded to the application, the intention is to execute the expert system as the primary thing to determine the distribution of data and the presence or absence of attack, and in case of presence of attack, to that of what kind throw the packet. Of these 'attack' packets created in the dataset, 30% of the data were used as input for verification, and 70% were used for training the various algorithms shown in Figure 2. After the expert system was used for attack detection, the output was compared Against the figure in Figure 2. The analysis results were made available to users in the last and third phases of the project using these various tools.

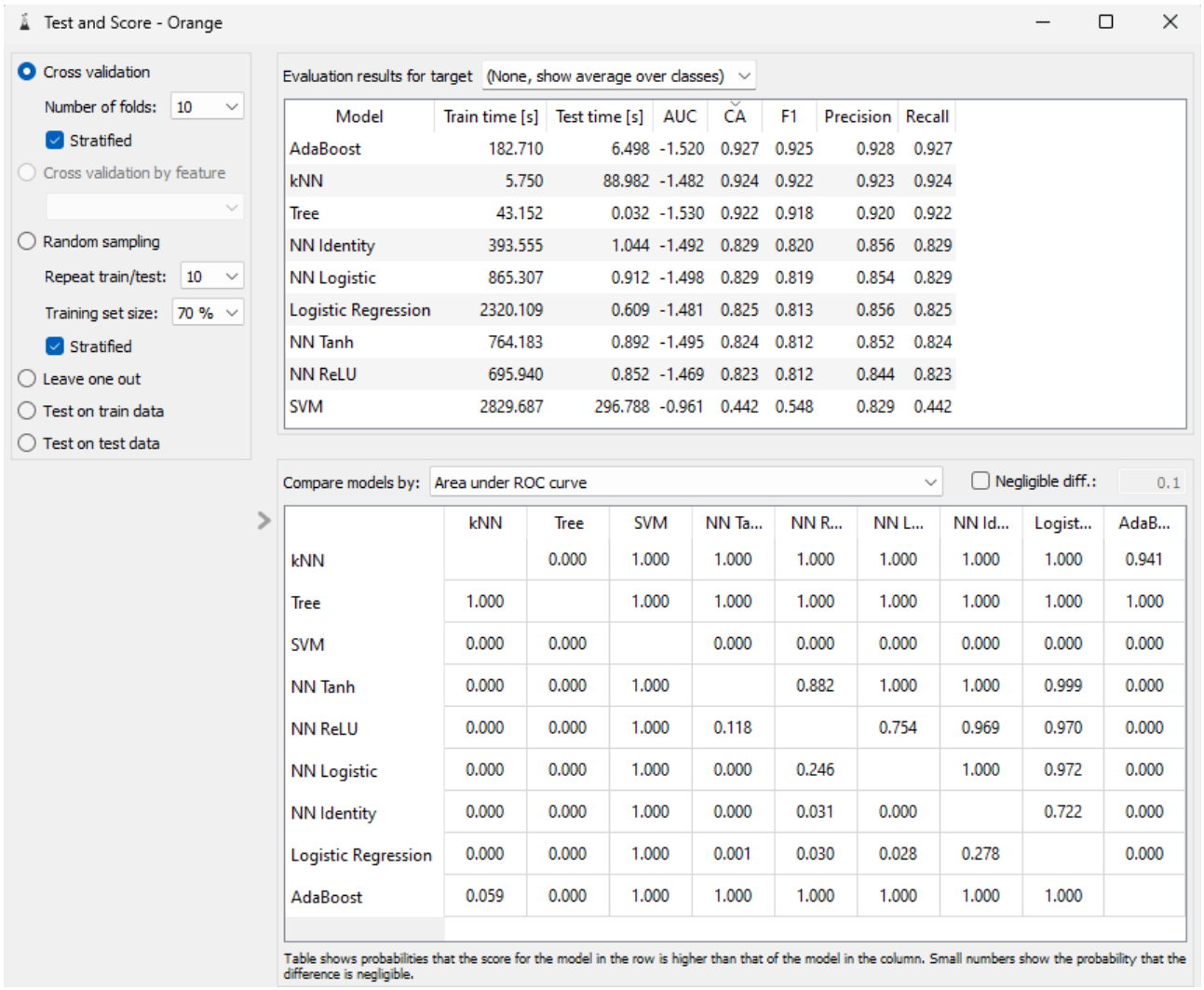


Figure 2. Model comparison table

The results in Figure 2 compare the essential values of different algorithms, such as accuracy, precision, recall, and F-score. The AdaBoost algorithm, which has the highest accuracy value of 92.7% among the compared results, was selected as the expert system, considering other essential values such as f-score, precision, and recursive precision.

In this way, using the AdaBoost algorithm in the expert system model provided higher accuracy and precision in intrusion detection and prevention.

5.1. Building and Training the Model

In the second phase of the intrusion detection analysis for IIoT, the AdaBoost algorithm was continued with the AdaBoost algorithm in the expert system since the AdaBoost model is the most successful in detecting attacks on the MQTT protocol used in IIoT through the expert system. The data was collected using natural systems and converted into a whitelist dataset. The created dataset was divided into 70% for training and 30% for validation. The Confusion Matrix of the Adaboost expert system model is shown in Figure 3. When the matrix in Figure 3 is analyzed, it is seen that the attack classification can be detected with high accuracy; in other words, there is no false-positive value.

Predicted

Actual	bruteforce	dos	flood	legitimate	malformed	slowite	Σ
bruteforce	5810	763	1	8	547	41	7170
dos	477	58687	1	4493	99	163	63920
flood	1	41	148	106	1	2	299
legitimate	0	1055	1	79977	4	0	81037
malformed	1575	518	0	49	3034	91	5267
slowite	677	375	0	618	95	2695	4460
Σ	8540	61439	151	85251	3780	2992	162153

Figure 3. Confusion matrix

Figure 4 shows the results of the Adaboost model. When Figure 4 is analyzed, it is seen that attacks on the MQTT protocol by the attacker at the same time as legitimate packets can be detected with 92.7% accuracy. This demonstrates the importance of continuous monitoring of network traffic in intrusion detection. This result is significant in ensuring the security and continuity of IIoT systems, which play an essential role in the communication control of critical infrastructures.

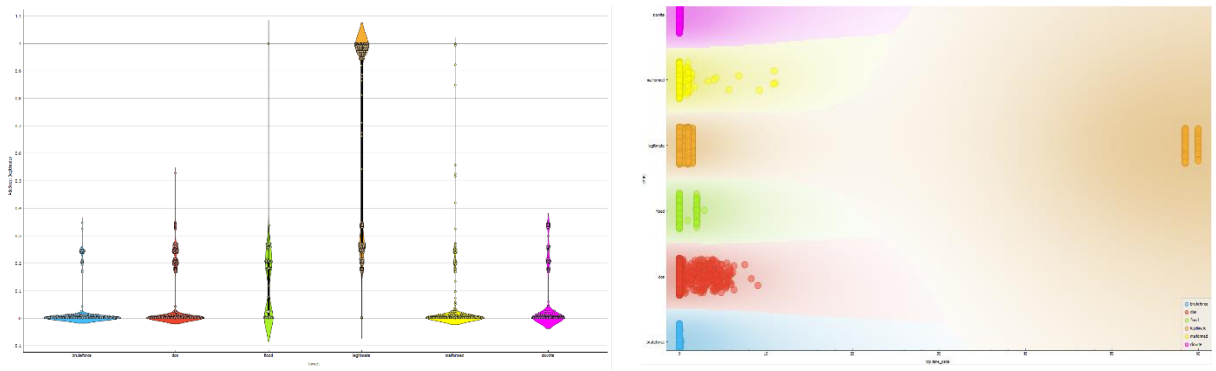


Figure 4. Adaboost expert system model results

The first graph shows the modeled results of various attacks against the MQTT protocol. The X-axis shows the different types of attacks, while the Y-axis shows the model's responses to the attacks. Each color represents a different type of attack:

Brute Force Attack (Blue): Brute force attacks usually result in a low response by the model. This indicates that the model has no difficulty detecting the attack as it is continuous and in a specific pattern.

DoS Attack (Red): DoS attacks cause a significant response in the model. The model must be able to detect this attack, as it generates heavy network traffic and consumes system resources.

Flood Attack (Orange): Flood attacks cause high and variable responses on the model. These attacks aim to degrade system performance by sending many requests to the network, and the model's ability to detect this attack is relatively high.

Legal Packages (Yellow): Legal packets represent the system's normal functioning and indicate that the model classifies these packets as usual.

Deconstructed Packets (Purple): Corrupted packets result in a high response in the model. This indicates that corrupted data is recognized as an anomaly in the system, and the model can detect these packets.

Slowite Attacks (Green): Slowite attacks cause a significant and wide range of responses in the model. This demonstrates the model's flexibility in detecting the attack in different ways.

The second graph shows the time delta (X-axis) and the model's responses (Y-axis) according to the attack types. This graph shows how the different types of attacks are distributed over time and the model's responses to these attacks.

Brute Force Attack (Blue): Brute force attacks are regularly distributed over time, and the model detected them. Attacks do not vary over time.

DoS Attack (Red): DoS attacks are concentrated at certain intervals along the time delta. The model consistently responds highly to these attacks over time.

Flood Attack (Orange): Flood attacks are widely distributed across the time delta, and the model reacts to them consistently well.

Legal Packets (Yellow): Legal packets are homogeneously distributed throughout the time delta, and the model reacts poorly. This represents the system's normal functioning.

Deconstructed Packets (Purple): Deconstructed packets are concentrated at certain intervals along the time delta, and the model responds highly to them.

Slowite Attacks (Green): Slowite attacks are widely distributed across the time delta, and the model reacts differently. This shows the dynamic nature of the attack and the model's resilience to it.

These graphs evaluated the model's ability to detect attacks and respond to attack types. They show the modeled results of various attacks against the MQTT protocol and how they are distributed over time.

6. Discussion

This study showed the capability of the expert system created to detect cyber attacks targeting MQTT-based IIoT systems. The Adaboost algorithm implemented in the study gave a high rate of accuracy, around 92.7%, thus improving on other algorithms. This is why Adaboost becomes very effective since it can aggregate weak classifiers into strong ones, hence detecting more advanced and complex types of attacks.

This study illustrated the application of comparison of specific elementary machine learning algorithms. For example, the results of decision trees or random trees were quite encouraging. Nevertheless, the fact that Adaboost is effective because it can obtain more accuracy with fewer resources was an essential consideration

in adopting this model. In particular, this model performed all-around even in infrequently focused attacks such as brute fair attacks and malformed packet attacks, which caused high success rates for this model.

One of the striking features of the system is the ability to function in real-time. In IIoT scenarios, managing real-time attacks without detection is critical since industrial activities will likely be disrupted. The expert system developed was able to identify the attacks even with high network usage and still do the work without any degradation in performance. This presents an advantage, especially regarding scaling and suitability for the industries.

Moreover, it is also pertinent that the system's structure allows it to learn continuously to meet any new threats that are likely to come in the future. This is because industrial Internet of Things IIoT systems work in a threat landscape that is always altering, and a static security approach may not be adequate in such cases. However, the expert system constructed within this investigation can adapt and learn new information and discover new attacks that have not been exhibited before. This enhances the system's security in the long run and is a better way of providing security in IIoT applications.

This work has, however, some limitations. First, the testbed would reproduce a small number of attack types. Actual attacks tend to be more complicated and varied, so applying the system to a more extensive dataset would be beneficial. Further, incorporating the system with other protocols could fill the remaining security holes within the IIoT system. In particular, integrating cryptographic measures with machine learning technologies will enhance protection.

Testing this expert system on other IoT and IIoT protocols will also be essential to ascertain its expected security level performance in future works. In addition, implementing and testing a wider range of attacks on the system and measuring how well the system manages these attacks will be invaluable research in the future. This research has successfully provided an efficient strategy for securing MQTT-based IIoT systems against various attacks. It has improved the existing body of knowledge in this area.

6.1. Comparison with a Wider Range of Attacks and Their Analysis

To effectively evaluate the robustness of the expert system developed for securing MQTT-based IIoT environments, it is essential to consider a comparative analysis encompassing a wider range of cyber-attacks beyond those initially tested, such as brute-force, DDoS, and malformed packet attacks. This section integrates insights from the literature to outline how the system could be benchmarked against more diverse and sophisticated threats and highlights the implications of such comparisons.

Advanced Persistent Threats (APTs): APTs represent prolonged and targeted cyber-attacks often orchestrated by well-resourced threat actors. Unlike the more immediate and high-visibility nature of DDoS or brute-force attacks, APTs are characterized by their stealth and persistence. They infiltrate networks, establish long-term footholds, and exfiltrate sensitive data over extended periods. Literature by Sharma and Bhushan [20] outlines a hybrid framework using Physically Unclonable Functions (PUFs) and machine learning to defend against such attacks in MQTT-based IoT systems. By extending the analysis to APTs, the expert system could be assessed for its capability to detect nuanced patterns associated with prolonged and low-frequency attacks, which are crucial in protecting industrial environments.

Reflection-Based Attacks (DRDoS): Distributed Reflection Denial of Service (DRDoS) attacks, as examined in studies like those by Kombate et al. [5], amplify malicious traffic by leveraging legitimate third-party servers. These attacks are challenging to identify due to their indirect nature. Integrating the detection of DRDoS within the expert system would require adaptive algorithms capable of recognizing subtle discrepancies in traffic flows and response patterns. Comparative analysis against such attacks would demonstrate the system's ability to maintain high accuracy in detecting traffic anomalies resulting from amplified, reflected sources.

Zero-Day Exploits: Zero-day vulnerabilities, highlighted in works by Deng [15] and Prajisha and Vasudevan [16], are previously unknown security flaws that attackers exploit before patches become available. These exploits pose significant risks to IIoT environments due to their unpredictability and potential to bypass conventional security measures. Incorporating detection models trained on evolving signatures and anomaly-based learning could enhance the expert system's resilience. Evaluating the system against simulated zero-day scenarios would showcase its potential to identify and mitigate unknown threats.

Multi-Vector Attacks: Complex, multi-vector attacks combine different techniques—e.g., combining DDoS with data exfiltration—to maximize disruption. Literature such as the study by Francis et al. [13] emphasizes how multi-layered detection systems, leveraging genetic algorithms and ensemble learning, can effectively counter such threats. The expert system's performance could be measured against these multi-faceted threats by implementing layered defensive strategies, including correlation-based detection and cross-protocol monitoring. This would illustrate its ability to adapt and respond to hybrid attack profiles.

Cryptographic and MITM Attacks: Research by Mahajan et al. [4] and Saqib and Moon [8] points to vulnerabilities in the MQTT protocol that can be exploited through cryptographic weaknesses or Man-in-the-Middle (MITM) attacks. Integrating encryption algorithms like AES and multi-factor authentication systems into the expert system could enhance its defensive capabilities. A comparative analysis of these attack types would provide valuable insights into the system's efficiency in maintaining data integrity and preventing unauthorized interception of communications.

6.2. Implications for Real-World Application

A comparative analysis involving a broader spectrum of attacks would underscore the expert system's comprehensive defensive capabilities and areas for improvement. The system could better withstand a dynamic threat landscape by incorporating multi-class machine learning models and cryptographic methods, as suggested by Liu et al. [7]. Additionally, continuous learning models that adapt to novel attack patterns—as discussed in the studies by Koprov et al. [9] and Zuhairi et al. [19]—would further enhance its long-term efficacy. This extended evaluation would validate the system's existing strengths and refine its ability to detect and respond to emerging threats, solidifying its utility in safeguarding MQTT-based IIoT infrastructures against a comprehensive range of cyber-attacks.

7. Conclusion

This study shows how efficient an expert system developed to secure IIoT systems is. The developed system protects against several cyber-attacks and can identify attacks such as brute force, DDoS, or corrupted packet attacks. It was noted that the expert system utilizing the AdaBoost algorithm managed attacks with precision as high as 92.7%. This high success rate was achieved due to Adaboost's ability to build more robust models by integrating several weak classifiers. This shows that the system is crucial in securing IIoT systems as it can identify even sophisticated types of attacks.

Even though the other machine learning algorithms employed in this work, Decision Trees and Random Forest, gave acceptable results, the high accuracy and low amount of resources used to implement the Adaboost algorithm were decisive in selecting this algorithm. Besides, the system's ability to function in real-time is a huge advantage for industrial processes such as IIOT, which highly depends on continuous and uninterrupted operations. The ability of the system to remain functional amid high levels of network traffic and still detect attacks is also of great importance in scaling and applying the system in industries.

The ability of the system to learn constantly means that it will always be ready for any new threat that may be experienced in the future. Since IIoT systems are being utilized in a dynamic threat landscape, more than static security mechanisms may be required. This paper highlights that the expert system engineered in this study can assimilate new information and discern unencountered types of attacks for the first time. Taking the time

to consider the duration enhances the system's sustainability in that the system provides long-term security measures for IIoT applications. There are also some limitations in this study. First, only a limited number of attack types could be reproduced using the testbed. Real-life attacks can be more complex and diverse; thus, testing the system using a large dataset would be ideal. It would also be helpful to integrate the system with other protocols in resource-constrained IIoT to cover other security gaps. In particular, it can enhance security by combining machine learning approaches and cryptographic methods.

Assessing this expert system with other IoT and IIoT protocols will be crucial to determining its security efficacy in future work. In addition, attempting more sophisticated attack scenarios and studying the system's behavior in the face of these attacks may prove valuable in guiding future research. To sum up, this research has offered the appropriate security mechanism for attack detection of MQTT-based IIoT systems and added great value to the body of knowledge in that area.

Author Contributions

The author read and approved the final version of the paper.

Conflicts of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] M. Al-Fayoumi, Q. A. Al-Haija, *Capturing low-rate DDoS attack based on MQTT protocol in software defined-IoT environment*, Array 19 (2023) 100316 10 pages.

[2] M. I. Fikriansyah, S. A. Karimah, F. Setiadi, *Detection of DDOS attacks in IIoT case using machine learning algorithms*, International Conference on Data Science and Its Applications (ICoDSA), 2024 IEEE, pp. 117–121.

[3] E. Shahri, P. Pedreiras, L. Almeida, *A scalable real-time SDN-based MQTT framework for industrial applications*, IEEE Open Journal of the Industrial Electronics Society 5 (2024) (2024) 215–235.

[4] R. A. Mahajan, R. G. Mahajan, M. Tatiya, U. H. Mandekar, M. Shahakar, Y. Patil, *Enhancing MQTT security in the internet of things with an enhanced symmetric algorithm*, Journal of Electrical Systems 20 (1s) (2024) 126–137.

[5] Y. Kombate, P. Houngue, O. Samuel, *Securing MQTT: unveiling vulnerabilities and innovating cyber range solutions*, Procedia Computer Science 241 (2024) 69–76.

[6] F. Buccafurri, V. De Angelis, R. Nardone, *Securing mqtt by blockchain-based otp authentication*, Sensor 20 (7) (2020) 2002 14 pages.

[7] Z. Liu, T. Liang, J. Lyu, D. Lang, *A security-enhanced scheme for MQTT protocol based on domestic cryptographic algorithm*, Computer Communications 221 (2024) (2024) 1–9.

[8] M. Saqib, A. H. Moon, *A novel lightweight multi-factor authentication scheme for MQTT-based IoT applications*, Microprocessors and Microsystems 110 (2024) (2024) 105088 20 pages.

[9] P. Koprov, X. Fang, B. Starly, *Machine identity authentication via unobservable fingerprinting signature: A functional data analysis approach for MQTT 5.0 protocol*, Journal of Manufacturing Systems 76 (2024) (2024) 59–74.

[10] A. Lohachab, *ECC-based inter-device authentication and authorization scheme using MQTT for IoT networks*, Journal of Information Security and Applications 46 (C) (2019) 1–12.

[11] I. Vaccari, G. Chiola, M. Aiello, M. Mongelli, E. Cambiaso, *MQTTset, a new dataset for machine learning techniques on MQTT*, Sensors 20 (22) (2020) Article Number 6578 17 pages.

[12] B. Mishra, A. Kertesz, *The use of MQTT in M2M and IoT systems: A survey*, IEEE Access 8 (2020) 201071–201086.

[13] G. T. Francis, A. Souri, N. İnanç, *A hybrid intrusion detection approach based on message queuing telemetry transport (MQTT) protocol in industrial internet of things*, Transactions on Emerging Telecommunications Technologies 35 (9) (2024) 5030 15 pages.

[14] C. Patel, N. Doshi, *A novel MQTT security framework in generic IoT model*, Procedia Computer Science 171 (2020) 1399–1408.

[15] X. Deng, *Internet of things (IoT) intrusion detection system (IDS) for home networks*, Doctoral Dissertation The George Washington University (2024) Washington.

[16] C. Prajisha, A. Vasudevan, *An efficient intrusion detection system for MQTT-IoT using enhanced chaotic salp swarm algorithm and LightGBM*, International Journal of Information Security 21 (6) (2022) 1263–1282.

[17] G. Karacayılmaz, H. Artuner, *A novel approach detection for IIoT attacks via artificial intelligence*, Cluster Computing 27 (2024) (2024) 10467–10485.

[18] S. Aymene, K. Lyes, A. Mondher, B. Ahmed, *Securing IIoT Against DDoS attacks: A stochastic approach*, Global Information Infrastructure and Networking Symposium (GIIS) (2024) 1–6.

[19] M. F. Zuhairi, S. M. Ali, Z. Shahid, M. M. Alam, M. M. Su'ud, *Real-time feature engineering for anomaly detection in IoT-based MQTT networks*, IEEE Access 12 (2024) (2024) 25700–25718.

[20] A. Sharma, K. Bhushan, *A hybrid approach based on PUF and ML to protect MQTT based IoT system from DDoS attacks*, Cluster Computing 27 (2024) (2024) 13809–13834.