A novel approach detection for IIoT attacks via artificial intelligence

Gökçe Karacayılmaz[1] · Harun Artuner[2]

Received: 29 December 2023 / Revised: 21 April 2024 / Accepted: 22 April 2024 / Published online: 6 May 2024

Abstract

The Industrial Internet of Things (IIoT) is a paradigm that enables the integration of cyber-physical systems in critical infrastructures, such as power grids, water distribution networks, and transportation systems. IIoT devices, such as sensors, actuators, and controllers, can provide various benefits, such as performance optimization, efficiency improvement, and remote management. However, these devices also pose new security risks and challenges, as they can be targeted by malicious actors to disrupt the normal operation of the infrastructures they are connected to or to cause physical damage or harm. Therefore, it is essential to develop effective and intelligent solutions to detect and prevent attacks on IIoT devices and to ensure the security and resilience of critical infrastructures. In this paper, we present a comprehensive analysis of the types and impacts of attacks on IIoT devices based on a literature review and a data analysis of real-world incidents. We classify the attacks into four categories: denial-of-service, data manipulation, device hijacking, and physical tampering. We also discuss the potential consequences of these attacks on the safety, reliability, and availability of critical infrastructures. We then propose an expert system that can detect and prevent attacks on IIoT devices using artificial intelligence techniques, such as rule-based reasoning, anomaly detection, and reinforcement learning. We describe the architecture and implementation of our system, which consists of three main components: a data collector, a data analyzer, and a data actuator. We also present a table that summarizes the main features and capabilities of our system compared to existing solutions. We evaluate the performance and effectiveness of our system on a testbed consisting of programmable logic controllers (PLCs) and IIoT protocols, such as Modbus and MQTT. We simulate various attacks on IIoT devices and measure the accuracy, latency, and overhead of our system. Our results show that our system can successfully detect and mitigate different types of attacks on IIoT devices with high accuracy and low latency and overhead. We also demonstrate that our system can enhance the security and resilience of critical infrastructures by preventing or minimizing the impacts of attacks on IIoT devices.

Keywords Artificial Intelligence · Attack Detection · Expert System · ICS Security · IIoT Security

1 Introduction

With the constantly changing and evolving technology, the number and diversity of technological devices used in smart cities and networks, factories, power plants, and other fields have increased, and so have the contributions they make to services. Therefore, today, many tasks that have been previously carried out by humans are now done more efficiently using information systems. Industrial Control Systems (ICS) are widely used in critical infrastructures such as smart factories, power grids, water treatment plants, etc. [1]. Supervisory Control and Data Acquisition (SCADA) systems are used to monitor and manage processes and control field equipment in these ICSs [2]. Programmable Logic Controller (PLC) devices, which are an important part of SCADA systems, have been increasingly connected to intranet networks and then to the internet in order to improve performance, efficiency, continuity, economy, remote control, and early fault detection. With the impact of Industry 4.0 and the development of technology, PLC devices, which are an important part of

✉ Gökçe Karacayılmaz
gkaracayilmaz@hacettepe.edu.tr

Harun Artuner
harun.artuner@gmail.com

[1] Forensic Sciences, Hacettepe University, Ankara, Turkey

[2] Computer Science Dept, Faculty of Engineering, Hacettepe University, Ankara, Turkey

critical infrastructure systems, have adapted to new technology, supporting digital connections and Internet of Things (IoT) protocols, and have become one of the most important components in the Industrial Internet of Things (IIoT). As the use of machines in production activities increases, so do machine-to-machine communication systems. The emergence of the Internet of Things has also made it possible to use the advantages and conveniences brought by these systems in industrial systems. These developments have led to the emergence of the IIoT concept. Over time, the need for security in SCADA systems, which were initially opened to the intranet and later to the internet, has increased even more with the inclusion of IoT systems in industrial systems. With the increased use of IoT systems in the industrial sector, productivity and speed have increased, and costs have decreased, making a significant contribution to human life. However, the IIoT devices and protocols included in critical infrastructure systems have also added new security vulnerabilities to these systems. These new devices and protocols have also made IIoT devices a new target for attackers in gaining access to SCADA systems and critical infrastructure systems. Therefore, in this study, an analysis of the attacks on IIoT systems is carried out, examining the vulnerabilities and the effects of attacks on the system. Then, solutions for detecting these attacks are proposed.

The main motivation and contribution of this work are to propose a novel expert system that combines continuous monitoring and attack detection using artificial intelligence algorithms for PLC devices integrated with IIoT devices. This work differs from the existing literature in the following aspects:

- It focuses on three types of cyberattacks that are highly relevant and challenging for ICS security: Man-in-the-Middle, Distributed Denial of Service, and Start-Stop attacks.
- It introduces a new feature, namely the “dup and retransmission” rate, to detect Man-in-the-Middle attacks more effectively than previous methods.
- It employs a hybrid approach that combines rule-based and machine learning-based techniques to detect attacks with high accuracy and low false positive rates.
- It evaluates the performance of the proposed system on a real-world embedded system that mimics an industrial process controlled by a PLC device.

Therefore, this work contributes to the advancement of cyber security research and practice for ICS and IIoT domains.

The significant contributions of the study are:

- The MitM, DDoS, and Start-Stop attacks were carried out on IoT/IIoT network traffic.
- Mirroring was performed for the analyses conducted to secure the network, thus avoiding imposing additional burdens on the operating systems.
- An expert system that combined continuous monitoring and attack detection through artificial intelligence was utilized in the attack detection phase.

During the detection phase, an examination was conducted to assess the influence of specific features on the efficacy of attack detection. Through the integration of the “dup and retransmission for MitM attack, continuous monitoring for DDoS attack, RTT and TTL for Start-Stop attack” features, a novel characteristic divergent from those employed in existing literature for flood attack identification, the expert system was able to achieve a markedly high success rate in detecting the attack.

In subsequent portions of the research, the second section scrutinized analogous studies within the realm of wireless communication and IoT. The third section detailed the creation of the embedded system specific to this study. The fourth section delineated the steps involved in vulnerability scanning, executing attacks, and detection mechanisms within the embedded system. The study culminated in the conclusion segment, wherein recommendations were proposed concerning the reinforcement of communication security.

2 Literature review

In this section, studies on the detection of cyber security threats that may arise as a result of the integration of PLC devices, an important component of ICS, with IIoT devices, and the detection of these threats using artificial intelligence (AI) algorithms are examined. PLC devices are defined as devices that consist of hardware and software components that operate based on a program in accordance with the data received from inputs/outputs and are expected to operate with minimal human interaction in Smart Factories, Smart Cities, and Critical infrastructures [3, 4].

IoT devices are defined as devices to which sensors with unique features such as remote sensing, communication, and low power consumption can be connected with limited resources (such as power and bandwidth), and IIoT is a subset of IoT and application of IoT devices in industrial processes. IIoT devices are described as systems that facilitate the connection of devices such as PLCs, RTUs, sensors, and actuators, which are components of Operational Technology (OT) systems in ICS [6].

The first cyber attack on ICS systems took place in 1903 when Italian radio pioneer Guglielmo Marconi’s long-distance wireless telegraph presentation was hacked using Morse code [7]. In ICS, updates to the components are not allowed by high-level managers due to high costs, labor, and time needs, and the focus is on the business framework [8]. As the use of IIoT-supported PLC devices becomes a part of our lives in smart factories, smart cities, and critical infrastructures, the damage that attackers can inflict on these systems, which affect human life, will become even more critical. Gönen et al. changed the data on the device by injecting incorrect data into the M241 PLC device and proposed the LiFi model as a solution [9]. Yilmaz and Gönen carried out a start-stop attack on S7-1200 PLC devices on a testbed using real devices and used the signature-based Snort IDS system for attack detection. It was seen that the attack detection system was static due to its signature-based structure [10]. Gueye et al. propose a neural network based method to detect cyber attacks on Modbus protocol used in IoT/IIoT devices. This method shows that an NN with an embedding function can be effectively used to model whether an attack has occurred on a device and the class of attacks that have occurred [11].

Jakovljevic and Nedeljkovic proposed a host-based IDS based on a semi-supervised CNN algorithm for the detection of cyber attacks on communication links in ICS. In their studies, they first used the Secure Water Treatment (SWaT) testbed data to design their system and then used the data set based on real data from the Electro-Pneumatic Positioning System (DisEPP) in the Manufacturing Automation Laboratory to use the designed IDS system [12]. Abdelaty et al. developed the “A Deep Learning Solution for Anomaly Detection in Industrial Control Systems (DAICS)” framework for detecting anomalies on actuators using SWaT and Water Distribution (WADI) data [13]. Charilaou et al. proposed a System for Operational Technology Attack Detection (SOTAD) using the Binary Logistic Regression algorithm using SWaT and WADI datasets to detect operational attacks on IIoT devices and PLC devices used in the field [14].

Other studies (2022) using the SWaT dataset aim to detect attacks on Industrial Control Systems using A class Neural Network supervised anomaly detection method [4, 15]. Mohammed et al. carried out a denial of service attack on 3 PLC systems using Modbus protocol and proposed a supervised XGBoost algorithm for attack detection [16]. Aydogan et al. simulated various attacks on RPL protocol-based IIoT systems and proposed a genetic programming-based IDS solution for attack detection [17]. Rahman and Hossain carried out various attacks on the 6G IT-OT testbed they developed and attempted to detect the attacks using various deep-learning methods. As a result of their studies, they stated that the innovations that IT-OT convergence will bring for 6G and the attacks carried out could be detected by deep learning methods [18]. Kim and Lee suggested a cyber attack detection system using various deep learning models on the CIC-IDS-2017 dataset to counter malware attacks on IIoT devices in smart factories [19]. Zhang et al. used a dataset prepared by the Mississippi State University Infrastructure Protection Center in 2014 to detect cyber attacks on IIoT devices and proposed a Graph Intrusion Detection (GID) framework for this purpose [20].

Khan et al. introduced a robust security model aimed at enhancing the protection of Industrial Internet of Things (IIoT) networks by deploying a deep learning-based Intrusion Detection System (IDS) for real-time detection of cyberattacks within Internet Industrial Control Systems (IICS). Their model, utilizing a Long Short-Term Memory (LSTM) autoencoder design, seeks to efficiently identify invasive activities in IICS networks. Experimental validation on the Gas Pipeline dataset and the UNSW-NB15 dataset demonstrated the proposed IDS’s effectiveness, achieving accuracy rates of 97.95 and 97.62%, respectively, thereby outperforming other leading methods in the field [21]. His other research, Khan et al. propose an innovative deep learning framework aimed at enhancing the detection and understanding of cyber threats within Industrial Internet of Things (IIoT) networks. The framework leverages an autoencoder-based detection mechanism, integrating convolutional and recurrent networks to discern and explain cyber threats effectively. By employing a two-step sliding window technique, this study not only advances the anomaly detection capabilities in IIoT networks but also emphasizes the framework’s ability to provide explanations for its predictions, facilitating a deeper understanding of the underlying reasons for detected threats. The empirical results underscore the framework’s robustness in identifying malicious events across various evaluation metrics, significantly outperforming contemporary methods and reinforcing its potential as a practical solution in real-world IIoT applications [22].

Radoglou-Grammatikis et al. investigate the cybersecurity vulnerabilities of low-voltage distribution systems within the smart grid, focusing on False Data Injection (FDI) cyberattacks facilitated through Man in The Middle (MiTM) actions. The study elaborates on two specific FDI attack scenarios: one targeting communications between a smart meter and an Active Distribution Management System (ADMS) and another targeting communications between a smart inverter and ADMS. These attacks potentially lead to devastating outcomes by affecting the operation of distribution transformers. To counteract these threats, the authors propose an Artificial Intelligence (AI)-based Intrusion Detection System (IDS) that demonstrates efficient detection and mitigation of such cyberattacks, thereby enhancing the security posture of smart grid infrastructure. The effectiveness of the proposed IDS is validated through experimental results, showcasing its capability to identify and respond to FDI attacks with a high degree of accuracy [23].

The proposed work by Khan et al. (2023) introduces a novel Intrusion Detection System (IDS) model named Federated-Simple Recurrent Units (Federated-SRUs) aimed at enhancing the security of IoT-augmented Industrial Control Systems (ICSs). This model leverages an improved architecture of Simple Recurrent Units (SRUs) to reduce computational costs and mitigate the issue of gradient vanishing commonly encountered in recurrent networks. By employing a federated learning approach, the Federated-SRUs IDS model facilitates data aggregation across multiple ICS networks in a privacy-preserving manner, allowing for the collaborative development of a comprehensive IDS model. The effectiveness of the Federated-SRUs IDS model is validated through extensive experiments on real-world gas pipeline-based ICS network data, demonstrating its capability to detect intrusions accurately in real time without compromising privacy and security. The experimental results confirm that the Federated-SRUs model surpasses existing state-of-the-art approaches, making it a promising solution for safeguarding IoT-based ICS networks against cyber threats [24]. Louati et al. propose a decentralized Multi-Agent Reinforcement Learning (MARL) based IDS for more effective intrusion detection in big data networks. They tested their proposed system on the NSL-KDD benchmark dataset and achieved an accuracy rate of 97.44% [25]. Nanjappan et al. proposed the DeepLG SecNet approach that leverages a combination of deep learning techniques including Long Short Term Memory (LSTM), Gated Secure Network (SecNet) and Crossover Chaos Game Optimization (CCGO) to harden IoT devices against unauthorized access and cyber attacks. The proposed method achieved 98.92% accuracy on various samples collected from BoT-IoT dataset and NSL-KDD dataset [26].

Chander and Kumar developed an improved pelican optimization model with ensemble voting based anomaly detection (EPOA-EVAD) approach to secure industrial procedures. In their work, they aimed to enhance anomaly detection capabilities in IIoT using EPOA-EVAD technique. Their proposed method provides an optimization model that combines techniques such as Synthetic Minority Over-sampling Technique (SMOTE) and ensemble voting classifier [27]. Alkhudaydi, Krichen, and Alghamdi address the growing challenge of cybersecurity attacks on the Internet of Things (IoT) by leveraging machine learning (ML) and deep learning (DL) algorithms to detect such attacks effectively. Their study involves a comprehensive evaluation of ten distinct ML models, including both single classifiers and ensemble classifiers, as well as four DL architectures, utilizing the Bot-IoT dataset for analysis. Notably, the application of the SMOTE to address data imbalance significantly enhanced model performance, with CatBoost and XGBoost classifiers achieving remarkable accuracy rates of 98.19 and 98.50%, respectively. This research underscores the potential of ML and DL techniques, in conjunction with data balancing strategies like SMOTE, to improve the detection of cybersecurity threats in IoT networks [28].

More recent research by Radoglou-Grammatikis et al. delve into the cybersecurity vulnerabilities inherent in the digitization of high-voltage electrical power and energy systems (EPES), particularly focusing on False Data Injection Attacks (FDIAs) that compromise EPES state estimation. Their research categorizes FDIAs into two distinct types: GPS Spoofing Attacks and IEEE C37.118 FDIAs, both targeting Phasor Measurement Unit (PMU) measurements within a high-voltage IEEE 9-Bus transmission grid emulation. Implementing an Artificial Intelligence (AI)-based Intrusion Detection System (IDS) demonstrates the system's ability to detect these cyberattacks effectively. The evaluation showcases the significant impact of FDIAs on the grid's operational integrity and validates the proposed IDS's efficacy in safeguarding against such threats [29].

Kelli, Radoglou-Grammatikis, Lagkas, Markakis, and Sarigiannidis delve into the vulnerabilities of the Distributed Network Protocol 3 (DNP3) widely used in smart grids, to enhance the cybersecurity posture of critical infrastructure. They identify and describe vulnerabilities intrinsic to DNP3 through the execution of eight cyberattack scenarios. Moreover, they introduce a novel risk assessment methodology that combines Attack Defence Trees (ADTs) with the Common Vulnerability Scoring System v3.1 (CVSS). This method quantifies the risk of cyberattacks on DNP3-enabled infrastructures, aiming to secure these critical systems by identifying potential threats and evaluating the severity of these attacks for more informed mitigation strategies [30]. Sarker, Khan, Abushark, and Alsolami (2022) present an extensive overview of IoT security, emphasizing the integration of machine and deep learning technologies to enhance the security of IoT systems against cyber threats. Their study delineates the significant role of artificial intelligence in developing a dynamic security framework capable of adapting to the evolving landscape of IoT vulnerabilities and attacks. By analyzing various machine learning and deep learning models, the authors advocate for an intelligent, data-driven approach to secure IoT devices and networks, highlighting future research directions to mitigate emerging security challenges within the IoT ecosystem [31].

Amponis et al. explore the vulnerabilities of the 5G core network, particularly focusing on the Packet Forwarding Control Protocol (PFCP) within the context of unauthorized Denial of Service (DoS) attacks. Their work introduces and tests a series of attacks aimed at disrupting established 5G tunnels, demonstrating the feasibility of such attacks without affecting subscribers' connectivity to the Next Generation Radio Access Network (NG-RAN), thereby complicating detection. Through the development and deployment of these attacks in a simulated environment, the authors highlight significant security flaws within the 5G core, especially concerning the PFCP protocol's handling of session control packets. This comprehensive study not only reveals potential threats to 5G network stability and security but also emphasizes the need for enhanced protective measures and protocols within the 5G architecture to mitigate these vulnerabilities effectively [32]. Kelli, Radoglou-Grammatikis, Sesis, Lagkas, Fountoukidis, Kafetzakis, Giannoulakis, and Sarigiannidis delve into the security vulnerabilities of the Distributed Network Protocol 3 (DNP3) used in Industrial Control Systems (ICS)/Supervisory Control and Data Acquisition (SCADA) systems. Their comprehensive study uncovers the inherent vulnerabilities within DNP3, leading to the execution of eight specific cyberattack scenarios. To counter these threats, they developed and demonstrated a Deep Neural Network (DNN)–based multi-model Intrusion Detection System (IDS), which was trained on an experimental network flow dataset comprising cyberattack scenarios. The IDS exhibited a remarkable accuracy rate of 99.0%, showcasing its efficacy in classifying DNP3 cyberattacks and enhancing the cybersecurity posture of ICS/SCADA systems [33].

Mladenov, Chobanov, Sarigiannidis, Radoglou-Grammatikis, Hristov, and Zlatev address the pressing issue of cybersecurity in the energy sector, particularly focusing on the vulnerabilities of smaller hydropower plants like the Leshnitsa facility in Bulgaria, which are integral to the decentralized energy grid. Recognizing the critical need for affordable and effective cyber-defense mechanisms, the SPEAR consortium developed a comprehensive security platform tailored for diverse energy sector actors, emphasizing the enhancement of cybersecurity measures for smaller entities without prior security systems. This platform, tested at the Leshnitsa hydropower plant, showcases a multi-component toolset designed for the real-time detection, signalization, and forensic analysis of cyber-attacks, aiming to bolster the resilience and operational security of these vital energy production sites [34]. Mohy-Eddie et al. carried out a network intrusion detection system (NIDS) to mitigate smart agriculture security vulnerabilities and evaluated their model using NF-Bot-IoT and NF-ToN-IoT datasets and achieved 99.25% accuracy [35]. Sivasakthi et al. delve into a robust learning approach named HybridRobustNet (HRN) for predicting and detecting hybrid attacks over IoT networks. HRN integrates various DL and ML algorithms to achieve improved detection accuracy and resilience against evolving hybrid attack patterns. Their multi-layer simultaneous deep reinforcement learning system (HRN) has an accuracy of 0.99977 [36].

In most of the above studies, artificial intelligence analyses were conducted on pre-existing datasets or simulation programs, and results were presented. While the use of simulation and pre-existing datasets can be useful for modeling and testing complex systems, the results should be considered in relation to the success rate of simulating real systems. Also, it should be noted that when using existing datasets, the relevance and up-to-dateness of the dataset and whether it reflects recent developments in ICS should be taken into account. Therefore, in this study, analyses were performed on a testbed infrastructure consisting of real IIoT systems, and artificial intelligence algorithms were used on the resulting datasets to aim for attack detection.

3 Testing infrastructure (testbed)

The testing infrastructure created for the study in order to carry out the attack and detection stages on Programmable Logic Controllers (PLCs), which are an important component of industrial control systems, is described in Fig. 1. When Fig. 1 is examined, it can be seen that the infrastructure is divided into three main sub-sections. The first section consists of remote-controlled and field equipment such as RTU, PLC, and IIoT devices, which are the control devices and sensors of the ICS. The second section is the human–machine interface (HMI), which is responsible for the management and operation of the field equipment. These two sections can connect via an intranet. However, as stated in the Introduction, for reasons of economy, early detection and intervention, and efficiency, IIoT devices are also connected to management units on the Cloud via the internet. Each unit has its own vulnerabilities and threats. When the topology is examined, it can be seen that the data unit processing is carried out by the process of packet mirroring in order to avoid disrupting continuity, which is the most important security component of industrial control systems. Additionally, the system and devices in which the testbed design is implemented are specified with the colors described in Fig. 1. IoT and IIoT devices are shown in green, legal servers and control systems are shown in blue, and the attacker is shown in red. The servers in a cloud environment are implemented in a virtual environment, while IoT and IIoT systems are implemented on real devices.

The analysis shown in Fig. 2 has been performed to create an attack detection model by carrying out the attack and detection stages. An artificial intelligence-based expert

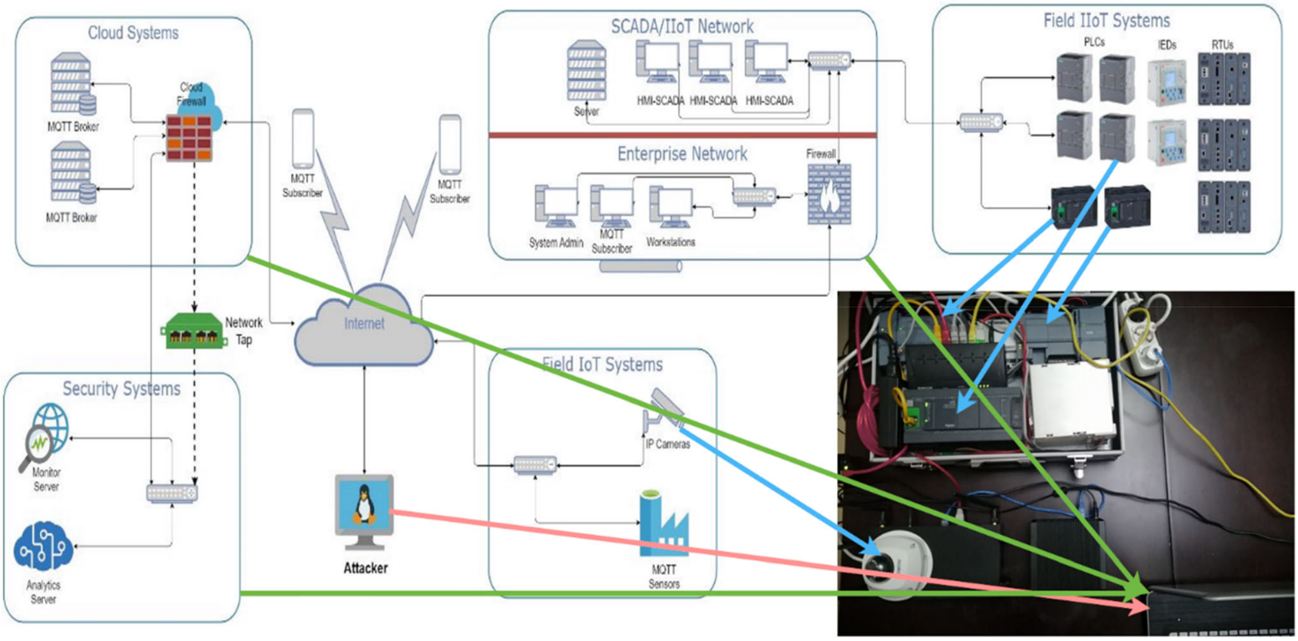


Fig. 1 Testbed Design

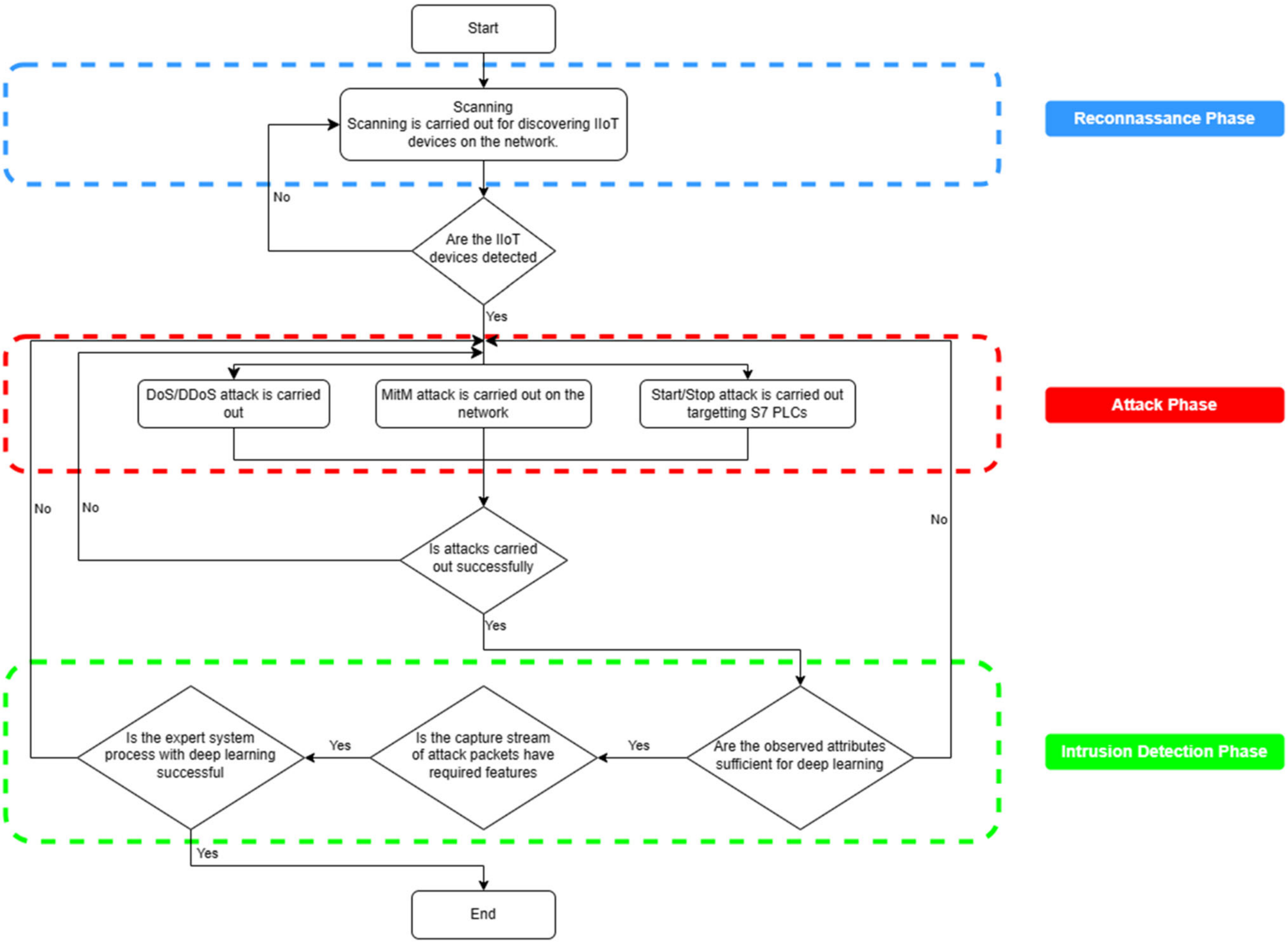

Start
Scanning
Scanning is carried out for discovering IIoT devices on the network.
Reconnassance Phase
No
Are the IIoT devices detected
Yes
DoS/DDoS attack is carried out
MitM attack is carried out on the network
Start/Stop attack is carried out targetting S7 PLCs
Attack Phase
No
No
No
Is attacks carried out successfully
Yes
Is the expert system process with deep learning successful
Yes
Is the capture stream of attack packets have required features
Yes
Are the observed attributes sufficient for deep learning
Intrusion Detection Phase
Yes
End

Fig. 2 Flowchart Diagram

system analysis method has been used to create the attack detection model. In the first stage of the tests, a network scan has been performed to determine whether there is a PLC device on the network. The purpose of PLC devices is to transfer sensor data received from IIoT devices, analyze them according to the programmed values, and manage motor systems. Because it is also used in many fields, such as smart factories, smart cities, and critical infrastructures, the start-stop attack is gaining importance. In this context, the devices on the network have been scanned using the open-source nmap tool. As a result of the scan, PLC devices/devices on the network have been determined based on the parameters used (brand, model, version number, etc.). In the second stage, the detection has been verified as correct or false positive by listening to the network traffic with a man-in-the-middle attack on the detected PLC device. In the next stage, the Hping3 tool has been used to perform a DDoS attack to distract the attention of network administrators and cybersecurity staff and perform a stealth attack before the actual attack. By sending packets through the tool, the network's service to legal users has been hindered. While network administrators and cybersecurity staff have been dealing with this attack, the final target attack has been carried out on the actively used S7-1200 PLC device by the starting-stopping attack on critical infrastructures. DDoS and start-stop attacks are very easy to detect by network administrators and cybersecurity staff when they occur because the effects of these active attacks on the system can be easily seen. However, man-in-the-middle attacks, which are passive attacks, are very difficult to detect, especially if continuous monitoring is not done, because they leave very little trace on the network.

All network traffic related to the process has been recorded using Wireshark, an open-source network traffic analysis software, which is to be used in the artificial intelligence-based expert system. The focus of the attack detection is the detection of the start-stop attack using a network packet that is very similar to the real HMI packet structure on a network where a stealth attack is performed with high accuracy using an artificial intelligence-based expert system. In the study, the attack can be detected with a success rate above the desired threshold value by using the expert system.

4 Architectural design of the proposed intrusion detection solution

The proposed intrusion detection solution consists of three main components: data acquisition, data processing, and data visualization. Figure 2 shows the overall architecture of the proposed solution.

Data acquisition is the process of collecting network traffic data from IIoT devices and PLCs using a network tap device. A network tap device is a hardware device that provides a way to access the data flowing across a network by creating a mirrored copy of the data without affecting the original data packets. The network tap device can be installed between the IIoT devices and PLCs or between the PLCs and the HMI to capture the network traffic data. The network tap device sends the captured data to a server where the data processing component is deployed.

Data processing is the core component of the proposed solution, where the artificial intelligence-based expert system is implemented. The data processing component receives the network traffic data from the data acquisition component and performs several steps to analyze the data and detect any attacks. The data processing component consists of four sub-components: packet filtering, feature extraction, classification, and alert generation.

Packet filtering is the process of filtering out irrelevant or redundant packets from the network traffic data. For example, packets that are not related to the communication between the IIoT devices and PLCs or between the PLCs and the HMI can be discarded. Packet filtering can reduce the size of the data and speed up the subsequent analysis steps.

Feature extraction is the process of extracting relevant features from the filtered packets that can be used to characterize the normal or abnormal behavior of the network. Features can be extracted from the packet headers or payloads, depending on the protocol and application used. For example, some features that can be extracted from the S7comm protocol packets are the function code, the parameter length, the data length, the error class, and the error code. Feature extraction can transform the raw network traffic data into a more structured and meaningful form that can be used for classification.

Classification is the process of applying the artificial intelligence-based expert system to classify network traffic data into normal or attack categories. The expert system uses a rule-based approach to identify the patterns and signatures of different types of attacks, such as network scan, MitM, DDoS, and start-stop attacks. The expert system can also use machine learning techniques to learn from historical or labeled data and improve its accuracy and adaptability. Classification can produce a binary or multi-class output that indicates the presence or absence of attacks and the type of attacks.

Alert generation is the process of generating alerts based on the output of the classification process. Alerts can contain information such as the timestamp, the source and destination IP addresses and ports, the protocol, the attack type, and the severity level of the attack. Alerts can be stored in a database or sent to the data visualization component for further analysis and action.

Data visualization is the process of presenting the alerts and the network traffic data in a graphical and interactive way to the cybersecurity staff or the system administrators. Data visualization can use dashboards, charts, graphs, maps, or tables to display the alerts and the network traffic data in a graphical and interactive way to the cybersecurity staff or the system administrators. Data visualization can use dashboards, charts, graphs, maps, or tables to display the information in an intuitive and easy-to-understand manner. Data visualization can also provide filtering, searching, sorting, and zooming functions to allow the users to explore the data and drill down into the details. Data visualization can help users monitor the network status, identify the attack sources and targets, assess the impact of the attacks, and take appropriate actions to mitigate the attacks.

5 Attack stages

In this section, the stages of the attacks carried out on the PLCs, which are the IIoT system components, on the testbed prepared within the scope of the study are explained. In this context, a network scan is first performed to detect PLC devices on the network, and then a man-in-the-middle (MitM) attack is performed to verify the detected devices. After obtaining basic information about the characteristics of the PLC devices, a Distributed Denial of Service (DDoS) attack is performed as a camouflage attack to distract the attention of the cybersecurity staff, who play a critical role in the system. Finally, the target attack of the study, the start-stop attack, is carried out in the last part of the attack analyses. During the attack analysis stage, the effects of all attacks on the system were primarily examined, and then the attack detection stage was entered.

The network scan is performed using the Nmap tool with the command “nmap -sS -A -vv -T4 192.168.1.*”, which performs a stealthy SYN scan, detects the service versions and operating systems of the hosts, increases the verbosity level, and uses the aggressive timing option. The network scan reveals that there are four PLC devices on the network, with IP addresses 192.168.1.101, 192.168.1.102, 192.168.1.103, and 192.168.1.104. The PLC devices are running the Schneider Electric firmware version V2.0.42.0 and have the Modbus TCP port 502 open for communication.

The next stage of the attack is to perform a man-in-the-middle (MitM) attack to intercept and modify the traffic between the PLC devices and the HMI. The MitM attack is carried out using the Ettercap tool, which can perform various techniques such as ARP poisoning, DNS spoofing, or DHCP spoofing to divert the traffic to the attacker’s machine. The command used to launch the Ettercap tool is “ettercap -G,” which opens the tool’s graphical user interface (GUI). The attacker then selects the PLC devices and the HMI as the targets and starts the ARP poisoning attack, which tricks the targets into thinking that the attacker’s machine is the gateway. As a result, the attacker can see and manipulate the packets exchanged between the PLC devices and the HMI, as shown in Fig. 4. The attacker can also use the Wireshark tool to capture and analyze the packets in more detail. The MitM attack allows the attacker to verify the presence and functionality of the PLC devices, as well as to learn about their configuration and settings.

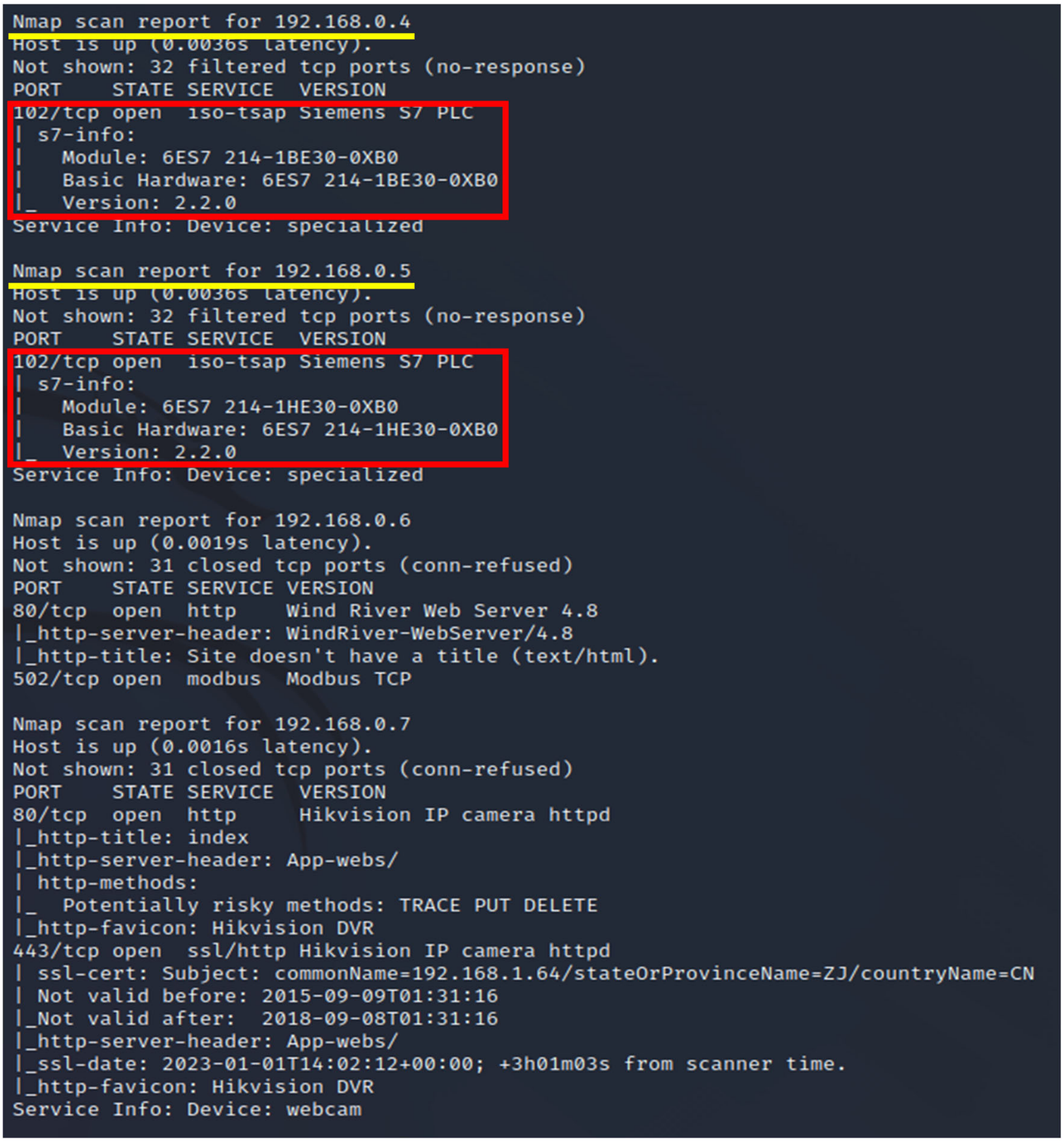
5.1 Attack analysis and effects on the system

In the first stage of the attack analysis, the Nmap network scanning tool is used to detect embedded systems on the network. As a result of the scan, IoT and IIoT devices on the network are identified, as shown in Fig. 3. As a result of the network scan, the IP address and information of the PLC devices to be targeted are determined.

5.2 MitM attack

In the second stage of the attacks on PLC devices, a man-in-the-middle attack, a passive attack type that is difficult to detect by network administrators if not specifically searched for, has been carried out. A man-in-the-middle attack is a type of attack where the malicious attacker enters between the two systems, communicating, listening to the communication, imitating both sides and accessing the information between them. The main goal of a MitM attack is to compromise the confidentiality of the communication or messages by secretly listening to the communication. This attack is usually carried out using address resolution protocol (ARP) security vulnerabilities. ARP is a protocol that helps match an IP address with the device’s MAC address in the local network [37].

The main goal of carrying out a MitM attack in the study is to confirm the systems that will be carried out an attack on and to collect the legal network packets required for a replay attack type called a start-stop attack. As a result of this attack, the data obtained about PLC with the Nmap application has been verified, and the packets required for the start-stop attack have been successfully collected from the network. In the scope of the study, IP information of PLCs obtained in the network scan has been used. In this way, ARP poisoning has been carried out on PLC devices, and important information shown in Fig. 4 has been obtained. With this information, vulnerabilities of these



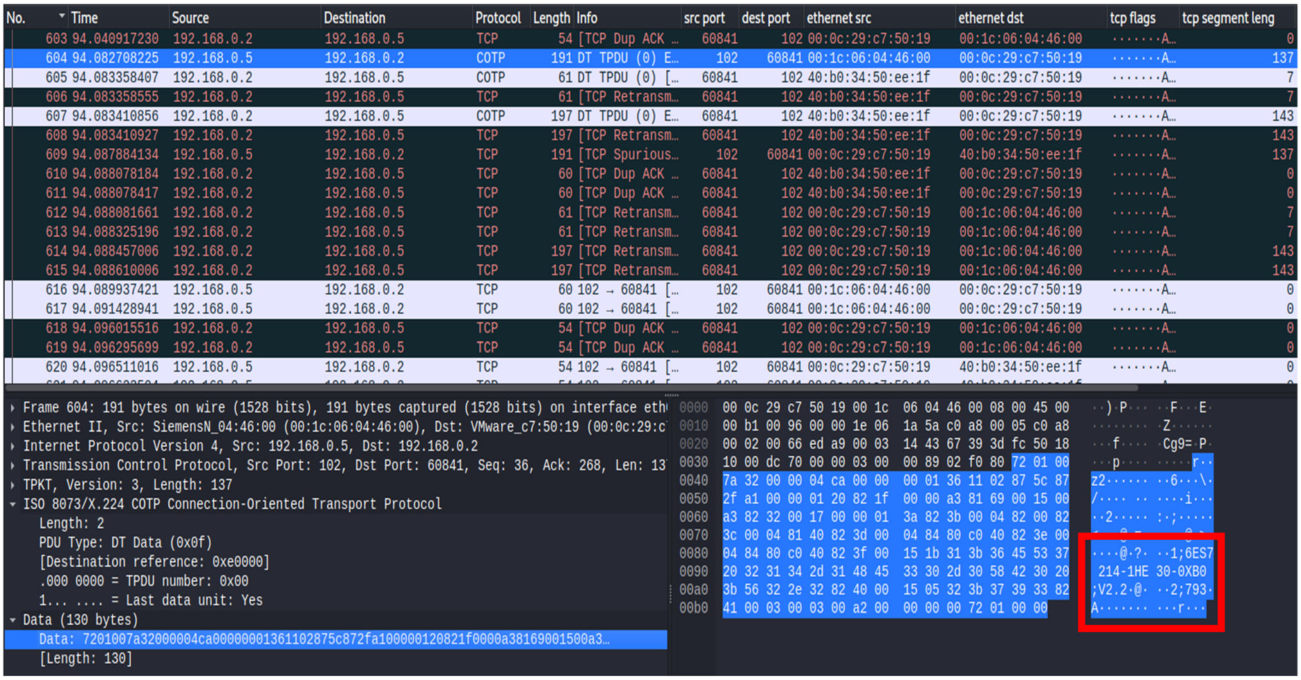
```
Nmap scan report for 192.168.0.4
Host is up (0.0036s latency).
Not shown: 32 filtered tcp ports (no-response)
PORT    STATE SERVICE  VERSION
102/tcp open  iso-tsap Siemens S7 PLC
| s7-info:
|   Module: 6ES7 214-1BE30-0XB0
|   Basic Hardware: 6ES7 214-1BE30-0XB0
|_  Version: 2.2.0
Service Info: Device: specialized

Nmap scan report for 192.168.0.5
Host is up (0.0036s latency).
Not shown: 32 filtered tcp ports (no-response)
PORT    STATE SERVICE  VERSION
102/tcp open  iso-tsap Siemens S7 PLC
| s7-info:
|   Module: 6ES7 214-1HE30-0XB0
|   Basic Hardware: 6ES7 214-1HE30-0XB0
|_  Version: 2.2.0
Service Info: Device: specialized

Nmap scan report for 192.168.0.6
Host is up (0.0019s latency).
Not shown: 31 closed tcp ports (conn-refused)
PORT    STATE SERVICE VERSION
80/tcp  open  http    Wind River Web Server 4.8
|_http-server-header: WindRiver-WebServer/4.8
|_http-title: Site doesn't have a title (text/html).
502/tcp open  modbus  Modbus TCP

Nmap scan report for 192.168.0.7
Host is up (0.0016s latency).
Not shown: 31 closed tcp ports (conn-refused)
PORT    STATE SERVICE  VERSION
80/tcp  open  http     Hikvision IP camera httpd
|_http-title: index
|_http-server-header: App-webs/
| http-methods:
|_  Potentially risky methods: TRACE PUT DELETE
|_http-favicon: Hikvision DVR
443/tcp open  ssl/http Hikvision IP camera httpd
| ssl-cert: Subject: commonName=192.168.1.64/stateOrProvinceName=ZJ/countryName=CN
| Not valid before: 2015-09-09T01:31:16
|_Not valid after:  2018-09-08T01:31:16
|_http-server-header: App-webs/
|_ssl-date: 2023-01-01T14:02:12+00:00; +3h01m03s from scanner time.
|_http-favicon: Hikvision DVR
Service Info: Device: webcam
```

Fig. 3 Network Scanning

devices and threats targeting these vulnerabilities have been determined.

5.3 DoS/DDoS attack

In the third stage of the attacks, a Distributed Denial of Service attack has been carried out on the PLC devices whose IP addresses have been identified. DDoS attacks are known as attacks aimed at blocking or disrupting a working service. The primary goal of DDoS attacks is to create more load than network, computer systems, and hardware resources such as bandwidth, memory, and disk space can handle, thus making the system unusable [38]. In this study, a DDoS attack has been used for the purpose of concealment for the target attack of the study, the start-stop attack. To determine the impact of the attack on the system, both packet access times and the number of network packets during the attack have been compared to the reference (the situation without the attack). The PLC device has been listening to the situation before the attack with ping (ICMP) packets. The response time of the target system to the ping packets before the attack has been observed to be around 1 ms. After this detection, the DDoS attack has been launched on the target system using bogon



No.	Time	Source	Destination	Protocol	Length	Info	src port	dest port	ethernet src	ethernet dst	tcp flags	tcp segment leng
603	94.040917230	192.168.0.2	192.168.0.5	TCP	54	[TCP Dup ACK …	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	0
604	94.082708225	192.168.0.5	192.168.0.2	COTP	191	DT TPDU (0) E…	102	60841	00:1c:06:04:46:00	00:0c:29:c7:50:19	·······A…	137
605	94.083358407	192.168.0.2	192.168.0.5	COTP	61	DT TPDU (0) […	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	7
606	94.083358555	192.168.0.2	192.168.0.5	TCP	61	[TCP Retransm…	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	7
607	94.083410856	192.168.0.2	192.168.0.5	COTP	197	DT TPDU (0) E…	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	143
608	94.083410927	192.168.0.2	192.168.0.5	TCP	197	[TCP Retransm…	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	143
609	94.087884134	192.168.0.5	192.168.0.2	TCP	191	[TCP Spurious…	102	60841	00:0c:29:c7:50:19	40:b0:34:50:ee:1f	·······A…	137
610	94.088078184	192.168.0.2	192.168.0.5	TCP	60	[TCP Dup ACK …	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	0
611	94.088078417	192.168.0.2	192.168.0.5	TCP	60	[TCP Dup ACK …	60841	102	40:b0:34:50:ee:1f	00:0c:29:c7:50:19	·······A…	0
612	94.088081661	192.168.0.2	192.168.0.5	TCP	61	[TCP Retransm…	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	7
613	94.088325196	192.168.0.2	192.168.0.5	TCP	61	[TCP Retransm…	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	7
614	94.088457006	192.168.0.2	192.168.0.5	TCP	197	[TCP Retransm…	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	143
615	94.088610006	192.168.0.2	192.168.0.5	TCP	197	[TCP Retransm…	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	143
616	94.089937421	192.168.0.5	192.168.0.2	TCP	60	102 → 60841 […	102	60841	00:1c:06:04:46:00	00:0c:29:c7:50:19	·······A…	0
617	94.091428941	192.168.0.5	192.168.0.2	TCP	60	102 → 60841 […	102	60841	00:1c:06:04:46:00	00:0c:29:c7:50:19	·······A…	0
618	94.096015516	192.168.0.2	192.168.0.5	TCP	54	[TCP Dup ACK …	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	0
619	94.096295699	192.168.0.2	192.168.0.5	TCP	54	[TCP Dup ACK …	60841	102	00:0c:29:c7:50:19	00:1c:06:04:46:00	·······A…	0
620	94.096511016	192.168.0.5	192.168.0.2	TCP	54	102 → 60841 […	102	60841	00:0c:29:c7:50:19	40:b0:34:50:ee:1f	·······A…	0

```
› Frame 604: 191 bytes on wire (1528 bits), 191 bytes captured (1528 bits) on interface eth
› Ethernet II, Src: SiemensN_04:46:00 (00:1c:06:04:46:00), Dst: VMware_c7:50:19 (00:0c:29:c
› Internet Protocol Version 4, Src: 192.168.0.5, Dst: 192.168.0.2
› Transmission Control Protocol, Src Port: 102, Dst Port: 60841, Seq: 36, Ack: 268, Len: 13
› TPKT, Version: 3, Length: 137
˅ ISO 8073/X.224 COTP Connection-Oriented Transport Protocol
    Length: 2
    PDU Type: DT Data (0x0f)
    [Destination reference: 0xe0000]
    .000 0000 = TPDU number: 0x00
    1... .... = Last data unit: Yes
˅ Data (130 bytes)
    Data: 7201007a32000004ca00000001361102875c872fa100000120821f0000a38169001500a3…
    [Length: 130]
```

```
0000  00 0c 29 c7 50 19 00 1c  06 04 46 00 08 00 45 00   ··)·P··· ··F···E·
0010  00 b1 00 96 00 00 1e 06  1a 5a c0 a8 00 05 c0 a8   ········ ·Z······
0020  00 02 00 66 ed a9 00 03  14 43 67 39 3d fc 50 18   ···f···· ·Cg9=·P·
0030  10 00 dc 70 00 00 03 00  00 89 02 f0 80 72 01 00   ···p···· ·····r··
0040  7a 32 00 00 04 ca 00 00  00 01 36 11 02 87 5c 87   z2······ ··6···\·
0050  2f a1 00 00 01 20 82 1f  00 00 a3 81 69 00 15 00   /···· ·· ····i···
0060  a3 82 32 00 17 00 00 01  3a 82 3b 00 04 82 00 82   ··2····· :·;·····
0070  3c 00 04 81 40 82 3d 00  04 84 80 c0 40 82 3e 00   <···@·=· ····@·>·
0080  04 84 80 c0 40 82 3f 00  15 1b 31 3b 36 45 53 37   ····@·?· ··1;6ES7
0090  20 32 31 34 2d 31 48 45  33 30 2d 30 58 42 30 20    214-1HE 30-0XB0
00a0  3b 56 32 2e 32 82 40 00  15 05 32 3b 37 39 33 82   ;V2.2·@· ··2;793·
00b0  41 00 03 00 03 00 a2 00  00 00 00 72 01 00 00      A······· ···r···
```

Fig. 4 Sniffed Network Packets

IPs with the hping3 test tool, and the ping times have been observed to be over 100ms. Figure 5 shows the distribution of packets collected for the denial attack according to time.

Additionally, packets captured by the open-source packet analyzer Wireshark have been collected during a Distributed Denial of Service (DDoS) attack to analyze the packets generated on the system. The attack and collected packets for analysis are depicted in Fig. 6.

The Distributed Denial of Service (DDoS) attack has been successfully completed. It has been observed that during the attack, the response time of the PLC device to packets has increased. After it has been seen that the attack has been successfully carried out, a start-stop attack has also been launched on the PLC device while the attack has been still ongoing.

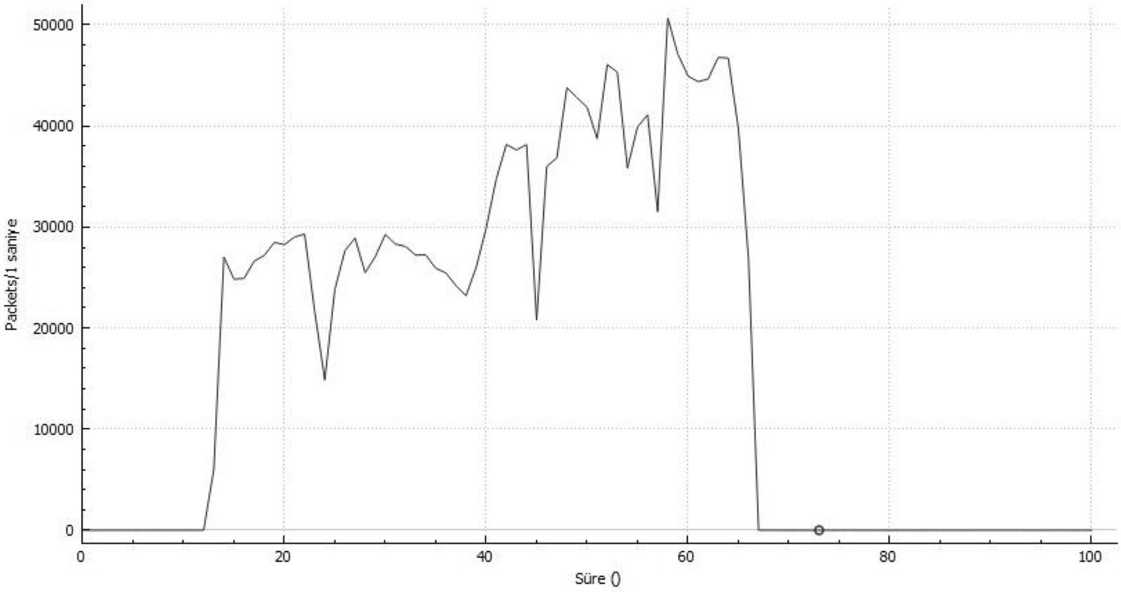
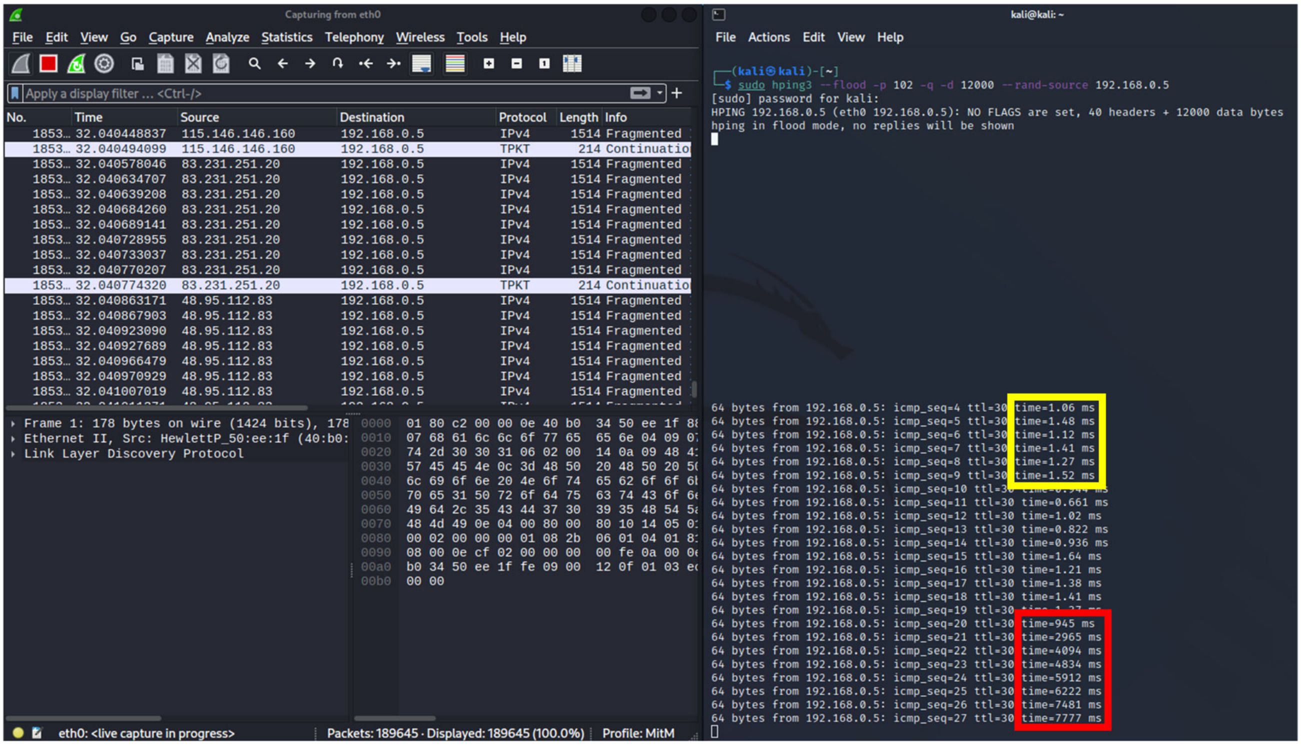


Fig. 5 DDoS Attack Packet Numbers over Time


Capturing from eth0
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter … <Ctrl-/>
kali@kali: ~
File Actions Edit View Help

```
No.      Time         Source           Destination   Protocol Length Info
  1853… 32.040448837  115.146.146.160  192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040494099  115.146.146.160  192.168.0.5   TPKT      214 Continuatio
  1853… 32.040578046  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040634707  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040639208  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040684260  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040689141  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040728955  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040733037  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040770207  83.231.251.20    192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040774320  83.231.251.20    192.168.0.5   TPKT      214 Continuatio
  1853… 32.040863171  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040867903  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040923090  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040927689  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040966479  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.040970929  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented
  1853… 32.041007019  48.95.112.83     192.168.0.5   IPv4     1514 Fragmented

› Frame 1: 178 bytes on wire (1424 bits), 178
› Ethernet II, Src: HewlettP_50:ee:1f (40:b0:
› Link Layer Discovery Protocol

0000  01 80 c2 00 00 0e 40 b0  34 50 ee 1f 8
0010  07 68 61 6c 6c 6f 77 65  65 6e 04 09 0
0020  74 2d 30 30 31 06 02 00  14 0a 09 48 4
0030  57 45 45 4e 0c 3d 48 50  20 48 50 20 5
0040  6c 69 6f 6e 20 4e 6f 74  65 62 6f 6f 6
0050  70 65 31 50 72 6f 64 75  63 74 43 6f 6
0060  49 64 2c 35 43 44 37 30  39 35 48 54 5
0070  48 4d 49 0e 04 00 80 00  80 10 14 05 0
0080  00 02 00 00 00 01 08 2b  06 01 04 01 8
0090  08 00 0e cf 02 00 00 00  00 fe 0a 00 0
00a0  b0 34 50 ee 1f fe 09 00  12 0f 01 03 e
00b0  00 00

eth0: <live capture in progress>   Packets: 189645 · Displayed: 189645 (100.0%)   Profile: MitM
```

```
┌──(kali㉿kali)-[~]
└─$ sudo hping3 --flood -p 102 -q -d 12000 --rand-source 192.168.0.5
[sudo] password for kali:
HPING 192.168.0.5 (eth0 192.168.0.5): NO FLAGS are set, 40 headers + 12000 data bytes
hping in flood mode, no replies will be shown

64 bytes from 192.168.0.5: icmp_seq=4 ttl=30 time=1.06 ms
64 bytes from 192.168.0.5: icmp_seq=5 ttl=30 time=1.48 ms
64 bytes from 192.168.0.5: icmp_seq=6 ttl=30 time=1.12 ms
64 bytes from 192.168.0.5: icmp_seq=7 ttl=30 time=1.41 ms
64 bytes from 192.168.0.5: icmp_seq=8 ttl=30 time=1.27 ms
64 bytes from 192.168.0.5: icmp_seq=9 ttl=30 time=1.52 ms
64 bytes from 192.168.0.5: icmp_seq=10 ttl=30 time=0.944 ms
64 bytes from 192.168.0.5: icmp_seq=11 ttl=30 time=0.661 ms
64 bytes from 192.168.0.5: icmp_seq=12 ttl=30 time=1.02 ms
64 bytes from 192.168.0.5: icmp_seq=13 ttl=30 time=0.822 ms
64 bytes from 192.168.0.5: icmp_seq=14 ttl=30 time=0.936 ms
64 bytes from 192.168.0.5: icmp_seq=15 ttl=30 time=1.64 ms
64 bytes from 192.168.0.5: icmp_seq=16 ttl=30 time=1.21 ms
64 bytes from 192.168.0.5: icmp_seq=17 ttl=30 time=1.38 ms
64 bytes from 192.168.0.5: icmp_seq=18 ttl=30 time=1.41 ms
64 bytes from 192.168.0.5: icmp_seq=19 ttl=30 time=1.37 ms
64 bytes from 192.168.0.5: icmp_seq=20 ttl=30 time=945 ms
64 bytes from 192.168.0.5: icmp_seq=21 ttl=30 time=2965 ms
64 bytes from 192.168.0.5: icmp_seq=22 ttl=30 time=4094 ms
64 bytes from 192.168.0.5: icmp_seq=23 ttl=30 time=4834 ms
64 bytes from 192.168.0.5: icmp_seq=24 ttl=30 time=5912 ms
64 bytes from 192.168.0.5: icmp_seq=25 ttl=30 time=6222 ms
64 bytes from 192.168.0.5: icmp_seq=26 ttl=30 time=7481 ms
64 bytes from 192.168.0.5: icmp_seq=27 ttl=30 time=7777 ms
```

Fig. 6 Network Traffic of DDoS Attack

5.4 Start-stop attack (targeted attack)

In the final attack of the study, under the attack analysis, a service denial masking attack has been carried out on the victim system, and the start-stop attack targeting the system's unintended repeated opening and closing is shown in Fig. 7. The Start-Stop attack is a replay attack type, in which the packets collected over the network by the attacker through MitM are analyzed and then sent to the PLC device again as desired. This attack involves collecting start and stop commands legally sent to the PLC device through the HMI interface and then modifying the protection bits in these packets to carry out the attack. This way, even if security experts/system administrators notice that the system is opening and closing while trying to find a solution to the service denial attack, they have tried to control it through the interface (Tia Portal for Siemens), but it has been seen that the system has returned to the attack loop.

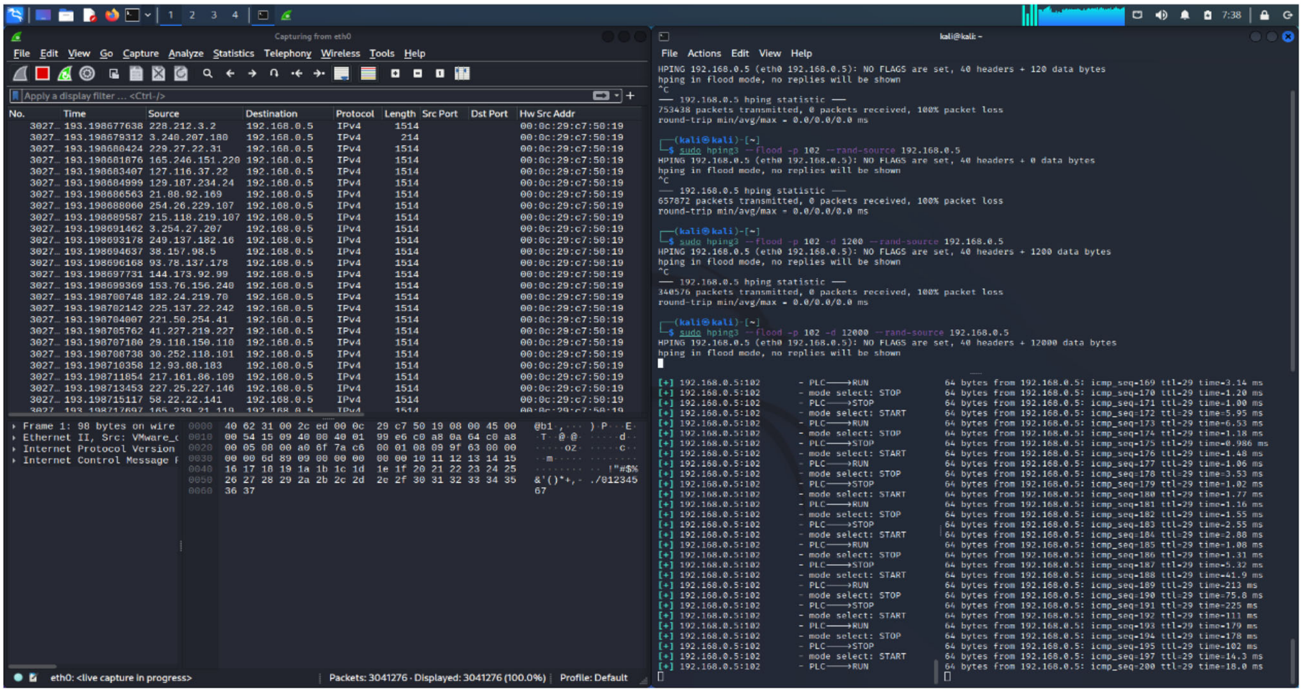
When only the Start-Stop attack is carried out on the system, the reference packet status in the system is seen in Fig. 8(a), and it is observed that the number of packets in the system increases when the Start-Stop attack is carried out under the service denial masking attack as shown in Fig. 8(b). When Fig. 8(b) is examined in detail, it is seen that there are interruptions in the transmission of attack packets caused by the DDoS attack during the attack. This is due to the system being unable to respond to DDoS packets during the shutdown, but no interruptions have been experienced during the.

Start-Stop attack.

In the study, the start-stop attack has been carried out with DDoS attack screening, so network administrators and cybersecurity personnel primarily deal with DDoS attacks, allowing the.

Start-stop attack to achieve its goal. The Start-Stop attack has been carried out on an S7-1200 PLC device, which is actively used in smart factories, smart cities, and critical infrastructure and also supports IIoT systems and protocols. The start-stop attack is the targeted attack of the study; it has been carried out to make the PLC open and close the system repeatedly in an unintended (not legal) way (100 times in this study). It has been observed that even if the operator tries to control the system via HMI during the attack, the system returns to the attack loop.

The study focuses on the detection of DDoS and Start-Stop attacks, which are considered active attacks, as their effects can be seen on the system and by experts/administrators. However, a MitM attack is a passive attack, and it does not have any effect on the system, making it difficult to detect. Therefore, features specific to each attack are selected to focus on detecting the attack through an expert system. The study focuses on the detection of the Start-Stop attack, which is the target attack, using an artificial


Capturing from eth0
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter … <Ctrl-/>
7:38
kali@kali: ~
File Actions Edit View Help

```
No.      Time            Source           Destination   Protocol  Length  Src Port  Dst Port  Hw Src Addr
  3027…  193.198677638  228.212.3.2      192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198679312  3.240.207.180    192.168.0.5   IPv4       214                        00:0c:29:c7:50:19
  3027…  193.198680424  229.27.22.31     192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198681876  165.246.151.220  192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198683407  127.116.37.22    192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198684999  129.187.234.24   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198686563  21.88.92.169     192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198688060  254.26.229.107   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198689587  215.118.219.107  192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198691462  3.254.27.207     192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198693178  249.137.182.16   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198694637  38.157.98.5      192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198696168  93.78.137.178    192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198697731  144.173.92.99    192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198699369  153.76.156.240   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198700748  182.24.219.70    192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198702142  225.137.22.242   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198704007  221.50.254.41    192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198705762  41.227.219.227   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198707180  29.118.150.118   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198708738  30.252.118.101   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198710358  12.93.88.183     192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198711854  217.161.86.109   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198713453  227.25.227.146   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198715117  58.22.22.141     192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19
  3027…  193.198717697  165.239.21.119   192.168.0.5   IPv4      1514                        00:0c:29:c7:50:19

› Frame 1: 98 bytes on wire       0000  40 62 31 00 2c ed 00 0c  29 c7 50 19 08 00 45 00   @b1·,··· )·P···E·
› Ethernet II, Src: VMware_c      0010  00 54 15 09 40 00 40 01  99 e6 c0 a8 0a 64 c0 a8   ·T··@·@· ·····d··
› Internet Protocol Version       0020  00 05 08 00 a0 6f 7a c6  00 01 08 09 9f 63 00 00   ·····oz· ·····c··
› Internet Control Message P      0030  00 00 6d 89 09 00 00 00  00 00 10 11 12 13 14 15   ··m····· ········
                                  0040  16 17 18 19 1a 1b 1c 1d  1e 1f 20 21 22 23 24 25   ········ ·· !"#$%
                                  0050  26 27 28 29 2a 2b 2c 2d  2e 2f 30 31 32 33 34 35   &'()*+,- ./012345
                                  0060  36 37                                              67

eth0: <live capture in progress>    Packets: 3041276 · Displayed: 3041276 (100.0%)    Profile: Default
```

```
HPING 192.168.0.5 (eth0 192.168.0.5): NO FLAGS are set, 40 headers + 120 data bytes
hping in flood mode, no replies will be shown
^C
—— 192.168.0.5 hping statistic ——
753438 packets transmitted, 0 packets received, 100% packet loss
round-trip min/avg/max = 0.0/0.0/0.0 ms

┌──(kali㉿kali)-[~]
└─$ sudo hping3 --flood -p 102 --rand-source 192.168.0.5
HPING 192.168.0.5 (eth0 192.168.0.5): NO FLAGS are set, 40 headers + 0 data bytes
hping in flood mode, no replies will be shown
^C
—— 192.168.0.5 hping statistic ——
657872 packets transmitted, 0 packets received, 100% packet loss
round-trip min/avg/max = 0.0/0.0/0.0 ms

┌──(kali㉿kali)-[~]
└─$ sudo hping3 --flood -p 102 -d 1200 --rand-source 192.168.0.5
HPING 192.168.0.5 (eth0 192.168.0.5): NO FLAGS are set, 40 headers + 1200 data bytes
hping in flood mode, no replies will be shown
^C
—— 192.168.0.5 hping statistic ——
340576 packets transmitted, 0 packets received, 100% packet loss
round-trip min/avg/max = 0.0/0.0/0.0 ms

┌──(kali㉿kali)-[~]
└─$ sudo hping3 --flood -p 102 -d 12000 --rand-source 192.168.0.5
HPING 192.168.0.5 (eth0 192.168.0.5): NO FLAGS are set, 40 headers + 12000 data bytes
hping in flood mode, no replies will be shown
█
```

```
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=169 ttl=29 time=3.14 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=170 ttl=29 time=1.20 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=171 ttl=29 time=1.00 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=172 ttl=29 time=5.95 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=173 ttl=29 time=6.53 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=174 ttl=29 time=1.18 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=175 ttl=29 time=0.986 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=176 ttl=29 time=1.48 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=177 ttl=29 time=1.06 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=178 ttl=29 time=3.53 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=179 ttl=29 time=1.02 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=180 ttl=29 time=1.77 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=181 ttl=29 time=1.16 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=182 ttl=29 time=1.55 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=183 ttl=29 time=2.55 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=184 ttl=29 time=2.88 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=185 ttl=29 time=1.08 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=186 ttl=29 time=1.31 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=187 ttl=29 time=5.32 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=188 ttl=29 time=41.9 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=189 ttl=29 time=213 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=190 ttl=29 time=75.8 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=191 ttl=29 time=225 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=192 ttl=29 time=111 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=193 ttl=29 time=179 ms
[+] 192.168.0.5:102     - mode select: STOP        64 bytes from 192.168.0.5: icmp_seq=194 ttl=29 time=178 ms
[+] 192.168.0.5:102     - PLC——→STOP               64 bytes from 192.168.0.5: icmp_seq=195 ttl=29 time=102 ms
[+] 192.168.0.5:102     - mode select: START       64 bytes from 192.168.0.5: icmp_seq=197 ttl=29 time=14.3 ms
[+] 192.168.0.5:102     - PLC——→RUN                64 bytes from 192.168.0.5: icmp_seq=200 ttl=29 time=18.0 ms
```

Fig. 7 Start-Stop Attack with Network Traffic during DDoS

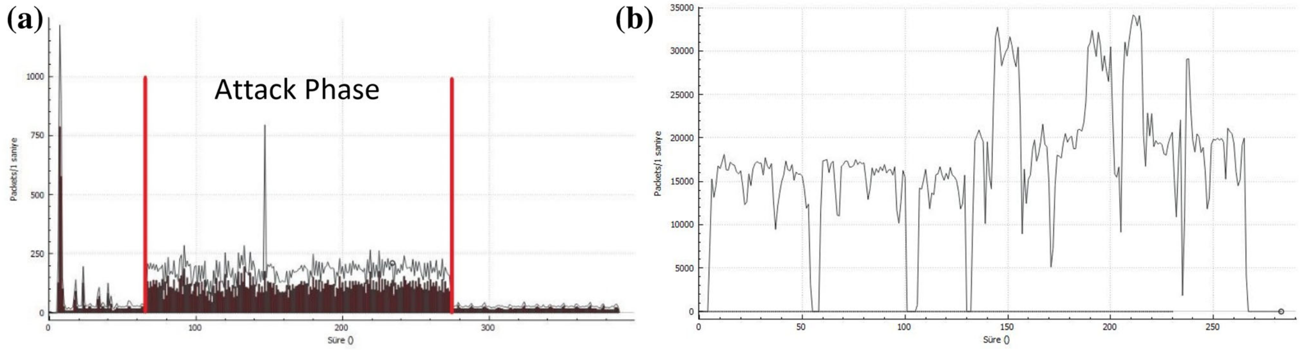


Fig. 8 **a** Start-Stop Attack. **b** Start-Stop Attack in DDoS Screening

intelligence-based expert system. Additionally, the detection of MitM and DDoS attacks are also carried out through the expert system.

6 Data analysis

In the scope of the study, data obtained from attack analyses have been evaluated in two different applications for expert system selection. Upon examination of Fig. 9, a structure of 3 (4, 10, 4) different layers has been determined, and the attack analyses have been examined based on this structure. In the examination, data obtained in the scope of attack analyses, when loaded into the application as a file, firstly determine the distribution of data and the most important dimension of whether or not there is an attack. Subsequently, data has been used for determining the features of the attack in terms of which type of attack (Attack classification) it is, and finally take, data samples to form the first phase of the expert system. Data from algorithms and remaining data are evaluated by testing in tenfold (second stage), as seen in Fig. 10 and the results of the analysis are visualized through various tools in the third stage.

In the expert system model, various traditional artificial intelligence algorithms and deep learning algorithms have been compared in Fig. 10. Among the compared results, the deep learning algorithm with the highest accuracy value

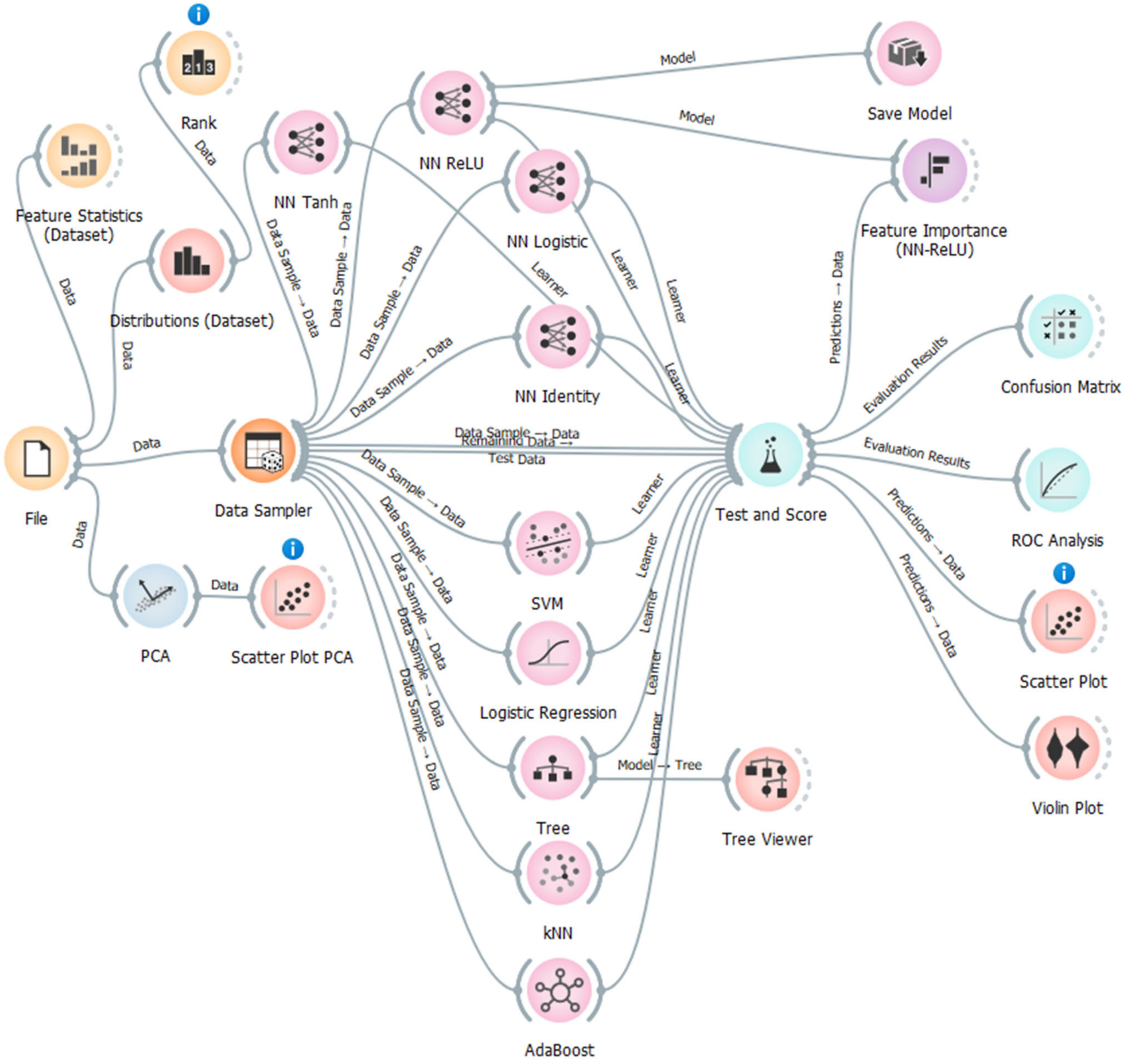


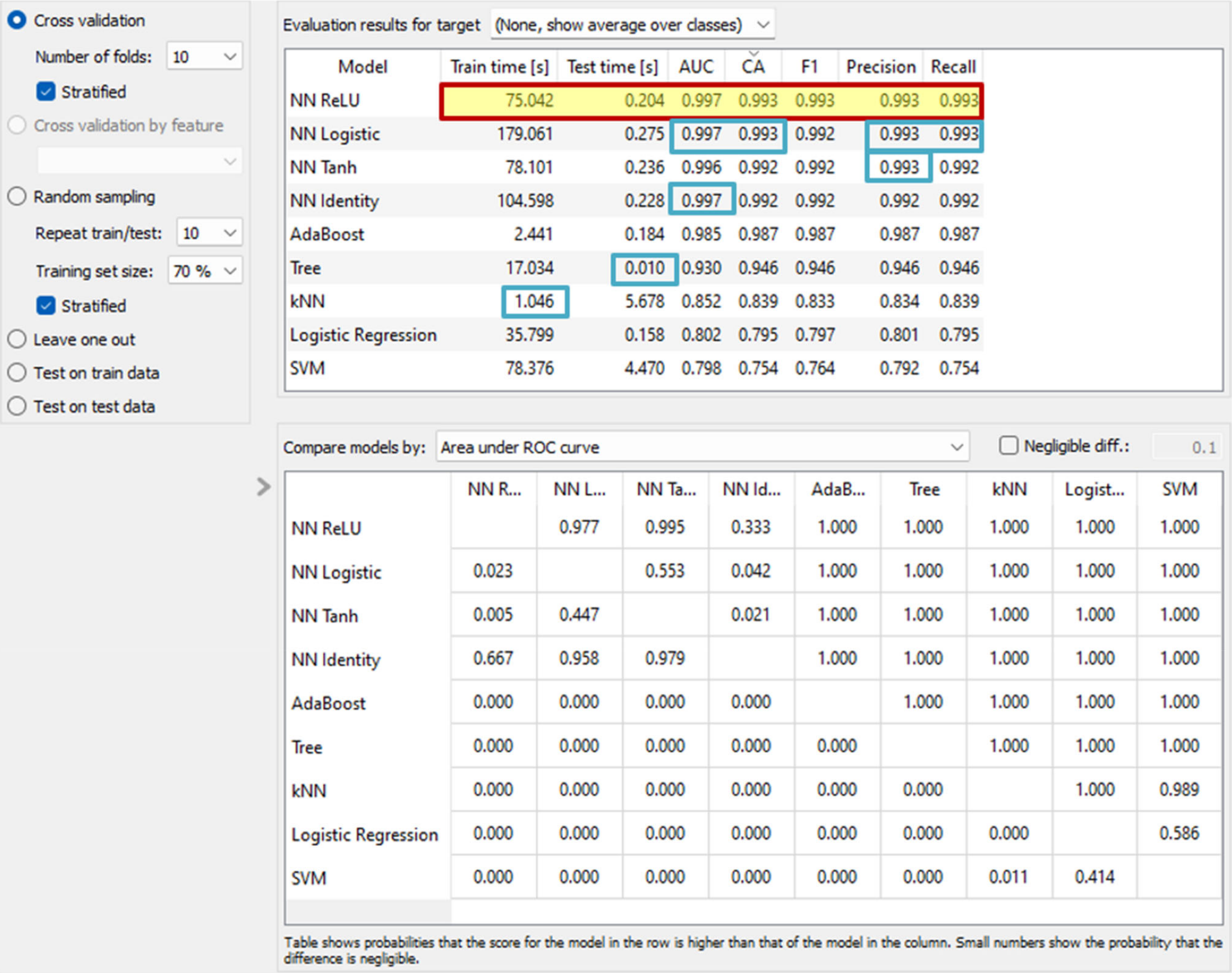
Fig. 9 Expert System Model

of 99.7% has been selected as the expert system with the Rectified Linear Unit (ReLU) for activation in the Neural Network model. Furthermore, these high accuracy rates are validated by the F1 score, precision, and recall values. Given the significant accuracy rates detected with the NN deep learning model, this study proceeded using the NN deep learning model as the chosen algorithm. The NN ReLU algorithm is seen to be more successful than other algorithms when all values are evaluated together, although NN Logistic and Tree algorithms also give results as successful as or better than NN ReLU when F1, CA, AUC, Precision, and Recall values are taken into account during testing time.

6.1 Neural network (NN) deep learning model

Artificial Neural Networks (ANNs) are mathematical models developed to analyze data using the working method of the human brain. Neural Network (NN) is one of these models. The basic unit of the model, as seen in Fig. 11, is an artificial neuron that mimics the human neuron. In humans, the output signals of neurons travel through synapse junctions with varying strengths and are then collected as input for the activation of a connected neuron.

A neural network can be simply defined as a neuron. This neuron takes inputs x1, x2, xn (and a + 1 bias value)



● Cross validation
Number of folds: 10
☑ Stratified
○ Cross validation by feature
○ Random sampling
Repeat train/test: 10
Training set size: 70 %
☑ Stratified
○ Leave one out
○ Test on train data
○ Test on test data

Evaluation results for target (None, show average over classes)

Model	Train time [s]	Test time [s]	AUC	CA	F1	Precision	Recall
NN ReLU	75.042	0.204	0.997	0.993	0.993	0.993	0.993
NN Logistic	179.061	0.275	0.997	0.993	0.992	0.993	0.993
NN Tanh	78.101	0.236	0.996	0.992	0.992	0.993	0.992
NN Identity	104.598	0.228	0.997	0.992	0.992	0.992	0.992
AdaBoost	2.441	0.184	0.985	0.987	0.987	0.987	0.987
Tree	17.034	0.010	0.930	0.946	0.946	0.946	0.946
kNN	1.046	5.678	0.852	0.839	0.833	0.834	0.839
Logistic Regression	35.799	0.158	0.802	0.795	0.797	0.801	0.795
SVM	78.376	4.470	0.798	0.754	0.764	0.792	0.754

Compare models by: Area under ROC curve
□ Negligible diff.: 0.1

	NN R...	NN L...	NN Ta...	NN Id...	AdaB...	Tree	kNN	Logist...	SVM
NN ReLU		0.977	0.995	0.333	1.000	1.000	1.000	1.000	1.000
NN Logistic	0.023		0.553	0.042	1.000	1.000	1.000	1.000	1.000
NN Tanh	0.005	0.447		0.021	1.000	1.000	1.000	1.000	1.000
NN Identity	0.667	0.958	0.979		1.000	1.000	1.000	1.000	1.000
AdaBoost	0.000	0.000	0.000	0.000		1.000	1.000	1.000	1.000
Tree	0.000	0.000	0.000	0.000	0.000		1.000	1.000	1.000
kNN	0.000	0.000	0.000	0.000	0.000	0.000		1.000	0.989
Logistic Regression	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.586
SVM	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.414	

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

Fig. 10 Comparison Table of Models

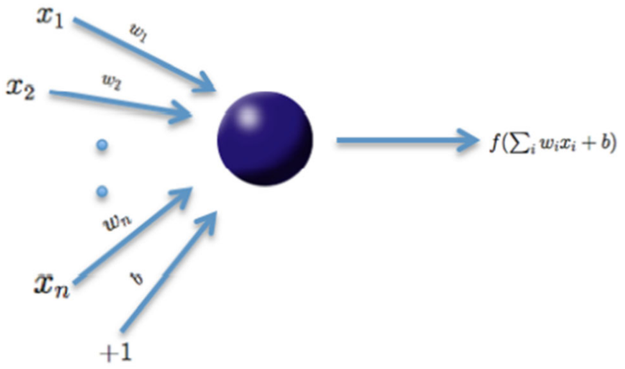


Fig. 11 Neuron Connection

and has an output that is a function f: R → R for $h_{W,b}(x) = f(W^T x) = f(\sum_{i=1}^{3} W_{ix_i} + b)$, which is known as the activation function. In the study, a multi-layer neural network algorithm with three nodes (4, 10, 4) is initialized with the ReLU [max = (0, x)] activation algorithm for the dataset, and the algorithm is run on the dataset. The model works on multiple layers of artificial neural networks that are connected to each other, as shown in Fig. 12, with neurons on each layer operating in parallel. The results of the deep learning model with multiple layers are described in detail in Sect. 5.2.

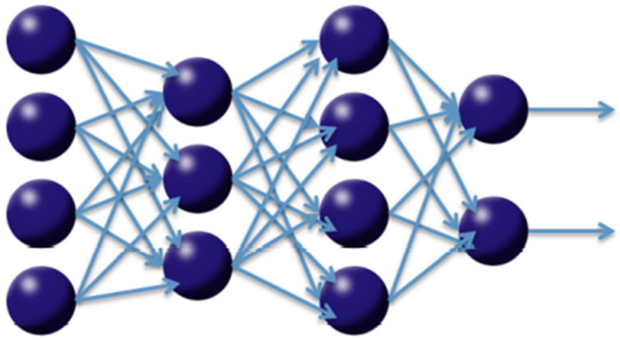


Fig. 12 Multi-layer Deep Learning Model

This study employs advanced deep-learning techniques to analyze data collected from mirrored networks. At the heart of these techniques are artificial neural networks (ANNs), which are organized in multiple layers. These layers work together to process data, allowing the system to identify patterns, make decisions, and learn independently. Essentially, this means the model can adapt and improve its accuracy over time without explicit programming for each task.

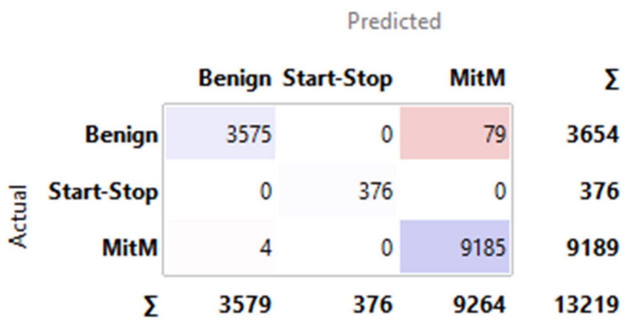
The process begins by feeding the collected data into a deep-learning processor. This processor is not one-size-fits-all; it utilizes various deep-learning methods that are selected based on the specific characteristics of the dataset at hand. By tailoring the approach to the data, the model can more effectively uncover the intricate structures and relationships within large and complex datasets. This adaptive, layered approach to data analysis enables the exploration of data in a way that's both deep and nuanced, leading to potentially groundbreaking insights in the study [39].

6.2 Creating and training the model

The study used an expert system artificial neural network model with the ReLU algorithm to detect start-stop packets during DDoS attack filtering. The Adam optimization algorithm has been used as the training model's solver. Data collected from real systems have been cleaned, labeled, and converted into a whitelist dataset for attack analysis and detection. The dataset has been then divided into 70% for training and 30% for testing. The most successful model, NN-ReLU, has been identified as the expert system model.

The Confusion Matrix for the expert system model is shown in Fig. 13. When Fig. 13 is examined, it provides information regarding the packets related to Man-in-the-Middle (MitM) and Start-Stop attacks, as evaluated by the Expert System. This information includes:

- Packets truly identified as attack packets.
- Benign network packets (non-attack) misclassified as attack packets.
- Packets incorrectly classified as attack packets when they are not.
- Start-stop attack packets misclassified as MitM attack packets.
- MitM attack packets misclassified as Start-Stop attack packets.



		Predicted			
		Benign	Start-Stop	MitM	Σ
Actual	Benign	3575	0	79	3654
	Start-Stop	0	376	0	376
	MitM	4	0	9185	9189
	Σ	3579	376	9264	13219

Fig. 13 Confusion Matrix

The Confusion Matrix is crucial for assessing the success of the expert system, as it allows us to focus on the packets that are genuinely attack packets and are correctly identified as such by the system. Therefore, the level of error in the system's reported success rate can be observed through this matrix. When evaluating the Confusion Matrix in this context, it becomes apparent that the numbers of false positives and false negatives are quite low compared to the actual attack packets in Fig. 13.

In the confusion matrix, the expert system has labeled 79 packets as MitM attack packets when they were actually normal network packets (Benign) and four packets as normal network packets when they were MitM attack packets.

The results of the NN-ReLU model over time are seen in Fig. 14, and upon examination of the results, it has been determined that legal packets have been being intercepted by MitM attacks, and start-stop attacks have been being made on the device identified as S7-PLC2 later. When examining the image on the left in Fig. 14, various types of attacks related to the traffic generated by legally sent packages (sent and received by HMI, PLC1, and PLC2) and the traffic produced by the attacker can be observed. In the analysis of the right side of the figure, the blue-marked legitimate network traffic packets, the red-labeled start-stop attack traffic packets, and the green-labeled attack packets related to the man-in-the-middle attack initiated by the attacker are depicted over the timeline by the expert system. Detailed information about these packets, including source IP address, port number, packet size, etc., can be viewed through the network analyzer. Consequently, continuous monitoring of the network by network/system experts or cybersecurity professionals, along with leveraging the expert system, enables the early detection of attacks and the implementation of preventive measures.

To evaluate the performance of the expert system, various metrics such as accuracy, true positive rate (TPR), false positive rate (FPR), and F1 score were calculated based on the confusion matrix. The results are shown in Table 1, where it can be seen that the expert system has achieved a high accuracy of 99.7% and a low FPR of 0.002, indicating that it can effectively distinguish between attack and normal packets. The TPR and F1 scores are also high, at 0.99 and 0.99, respectively, showing that the expert

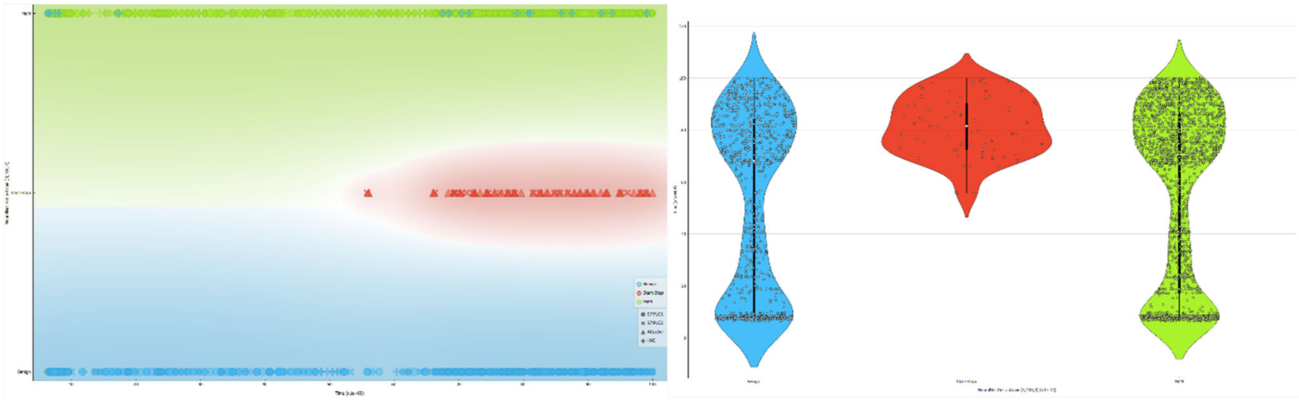


Fig. 14 NN-ReLU Expert System Model Results

Table 1 Performance metrics of the expert system

Model	Train time (s)	Test time (s)	AUC	CA	F1	Precision	Recall
NN relu	75.042	0.204	0.997	0.993	0.993	0.993	0.993
NN logistic	179.061	0.275	0.997	0.993	0.992	0.993	0.993
NN tanh	78.101	0.236	0.996	0.992	0.992	0.993	0.992
NN identity	104.598	0.228	0.997	0.992	0.992	0.992	0.992
Adaboost	2.441	0.184	0.985	0.987	0.987	0.987	0.987
Tree	17.034	0.010	0.930	0.946	0.946	0.946	0.946
KNN	1.046	5.678	0.852	0.839	0.833	0.834	0.839
Logistic regression	35.799	0.158	0.802	0.795	0.797	0.801	0.795
SVM	78.376	4.470	0.798	0.754	0.764	0.792	0.754

system can correctly identify most of the attack packets and minimize the false negatives.

To compare the proposed method with other AI methods, the same dataset has been used to train and test different models, such as Decision Tree, Adaboost, Support Vector Machine, and K-Nearest Neighbor. The comparison results are shown in Fig. 10, where it can be seen that the proposed NN-ReLU model outperforms the other models in terms of accuracy and F1 score. The proposed model also has the lowest FPR among the compared models, indicating that it is more robust against false alarms. Therefore, the proposed NN-ReLU model is the most suitable model for detecting attacks on IIoT systems in real-time.

7 Discussion

The analysis of the study aimed to first identify vulnerabilities in IIoT systems, then observe the effects of exploiting these vulnerabilities on the system and finally use data science to detect attacks in.

Real-time. In the first stage, IoT and IIoT devices have been detected during a scan of the network, and IIoT devices have been chosen as the target systems for the second stage due to the study's main objective. In the second stage, three different attacks targeted PLC devices. The first attack, the MitM attack, confirmed the information obtained during the scanning stage, and important information about the target system (brand, model, etc.) has been obtained through packet analysis. The second attack, the DDoS attack, has been used to block the system for security analysts and system administrators. The effects of the attack on the system have been observed through various components such as Time to Live (TTL), flags, Round Trip Time (RTT), and network traffic. The third attack, the start-stop attack, has been used to turn off and on IIoT systems unauthorized and in a certain loop (100 in the study). Even though the attack has been limited in its access to the system through the HMI, it has been unable to stop the loop and escape the effects of the attack.

In the final stage of the analysis, the data obtained from the attack analysis and the normal (reference) traffic data of the IIoT network without attack have been analyzed using data mining applications. The detection of specific features for each attack, the training of the system, and the validation of the test data resulted in the successful detection of attacks with a 99.7% accuracy rate using the NN deep learning algorithm.

8 Limitations

In this study, several limitations have been identified that could affect the generalizability and effectiveness of our proposed AI-based detection system for IIoT attacks. Firstly, the system's training and evaluation were confined to a particular set of data simulating industrial environments, which might not encompass the variability encountered in different industrial settings or emerging attack vectors. Additionally, while the system is optimized for accuracy and low latency within our testbed, its adaptability to ever-evolving cyber threats and scalability across more complex or larger network architectures requires further investigation.

Another significant limitation is the dependency on data quality. The AI model's performance is highly reliant on the comprehensiveness and integrity of the input data; thus, any shortcomings in the dataset could skew detection capabilities. Moreover, the system's real-time processing capabilities could be hindered under conditions of extreme data flow, potentially delaying the detection of attacks. Lastly, the computational and infrastructural demands for deploying and maintaining such a system might be prohibitive for smaller organizations with limited IT resources.

9 Future work

Given the identified limitations, the future work will focus on several key areas to enhance the robustness and applicability of the detection system. An immediate area of focus will be on improving the generalization of the model to ensure consistent performance across a variety of industrial sectors and under different cyber-attack scenarios. This will involve enriching the training datasets with a broader array of attack types and IIoT environments.

Moreover, optimizing the system architecture to manage larger volumes of data with minimal latency is crucial for supporting real-time detection capabilities in industrial contexts. Integrating the AI-based system with other cybersecurity frameworks could also provide a more comprehensive defense mechanism, offering robust protection against a wider spectrum of threats.

Exploring cutting-edge machine learning and deep learning strategies will be pivotal in reducing false positives and enhancing the accuracy of attack detection. Techniques such as federated learning could be particularly beneficial for deploying the system across multiple IIoT sites while preserving data privacy. Finally, developing cost-effective strategies that reduce both computational and economic burdens will make the system more accessible to a broader range of users, including small and medium-sized enterprises.

These future directions not only aim to address the current limitations but also anticipate adapting to the dynamic landscape of cybersecurity threats faced by industrial systems.

10 Conclusion

Information technologies have begun to be widely used in every aspect of our lives, from education to healthcare and from e-government transactions to social life. Cybersecurity is one of the most important dimensions of these infrastructures, and it provides important values such as efficiency, speed, and economy. The use of IoT and IIoT systems in Industry 4.0 and subsequent stages has opened the door to their effective use; however, it has also made the cybersecurity analysis of these systems very important due to the specific vulnerabilities that come with these technologies. Therefore, in this study, attacks that can be carried out on IoT and IIoT systems have been analyzed. In the attack analysis section of the study, the effects of the attacks on the system have been clearly detected through continuous network monitoring. Subsequently, a neural network model-based expert system model has been proposed to detect these attacks. This model is designed to help make IoT and IIoT systems more secure by collecting network packets through mirroring in ICS processes and analyzing and learning from the data. This method also helps to prevent the interruption of the continuity of ICS, which is the most important cybersecurity pillar, due to security analysis. A multi-layered neural network has been used to analyze the data collected from the network to detect attacks. This method helped to make the data used in industrial processes where IoT and IIoT systems are used more secure. The study also presents a detection analysis of these attacks, in which the authors propose the use of artificial intelligence algorithms to detect and prevent such attacks in a timely manner. In the attack detection analysis section, it has been observed that attacks have been detected with a 99.7% accuracy rate through artificial intelligence algorithms by specifically extracting features for each attack.

This study aims to ensure the continuity of the systems even in the case of similar attacks, minimizing the damage to the system. The authors also note that the methods proposed in this study can be applied to other brand and model IIoT systems as well.

Authors contributions Gökçe KARACAYILMAZ carried out the experiment. Gökçe KARACAYILMAZ and Harun ARTUNER

verified the analytical methods. Gökçe KARACAYILMAZ wrote the manuscript with support from Harun ARTUNER. All authors discussed the results and contributed to the final manuscript.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). The authors have not disclosed any funding.

Data availability If there are any requests, we can share our dataset.

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical approval Not applicable.

References

1. Kravchik, M., Shabtai, A.: "Detecting Cyber Attacks in Industrial Control Systems Using Convolutional Neural Networks", ser. CPS-SPC '18, pp. 72–83. Association for Computing Machinery, New York, NY, USA (2018)
2. Ayas, S., Ayas, M.S.: A modified densenet approach with nearmiss for anomaly detection in industrial control systems. Multimed. Tools. Appl. **81**(16), 22573–22586 (2021)
3. López-Morales E, Rubio-Medrano C, Doupé A, Shoshitaishvili Y, Wang R, Bao T, Ahn GJ (2020, October). HoneyPLC: a next-generation honeypot for industrial control systems. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security* (pp. 279–291).
4. Boateng EA (2021) Anomaly detection for industrial control systems based on neural networks with one-class objective function. *Proceedings of Student Research and Creative Inquiry Day*, *5*.
5. Kankanhalli, A., Charalabidis, Y., Mellouli, S.: IoT and AI for smart government: a research agenda. Gov. Inf. Q. **36**(2), 304–309 (2019)
6. Hansong, Xu., Wei, Yu., Griffith, D., Golmie, N.: A survey on industrial internet of things: a cyber-physical systems perspective. IEEE Access **6**(2018), 78238–78259 (2018)
7. Hemsley, K.E., Fisher, E.: History of industrial control system cyber incidents (No. INL/CON-18-44411-Rev002). Idaho National Lab.(INL), Idaho Falls, ID, United States (2018)
8. Ibarra J, Butt UJ, Do A, Jahankhani H, Jamal A (2019, January) Ransomware impact to SCADA systems and its scope to critical infrastructure. In *2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3)* (pp. 1–12). IEEE.
9. Gönen, S., Sayan, H.H., Yılmaz, E.N., Üstünsoy, F., Karacayılmaz, G.: False data injection attacks and the insider threat in smart systems. Comput. Secur. **97**, 101955 (2020)
10. Yılmaz, E.N., Gönen, S.: Attack detection/prevention system against cyber attack in industrial control systems. Comput. Secur. **77**, 94–105 (2018)
11. Gueye, T., Wang, Y., Rehman, M., Mushtaq, R.T., Zahoor, S.: A novel method to detect cyber-attacks in IoT/IIoT devices on the modbus protocol using deep learning. Clust. Comput. **26**(5), 2947–2973 (2023)
12. Nedeljkovic, D., Jakovljevic, Z.: CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. Comput. Secur. **114**, 102585 (2022)
13. Abdelaty, M., Doriguzzi-Corin, R., Siracusa, D.: DAICS: a deep learning solution for anomaly detection in industrial control systems. IEEE Trans. Emerg. Top. Comput. **10**(2), 1117–1129 (2021)
14. Charilaou C, Ioannou CI, Vassiliou V (2022, June) System for operational technology attack detection in industrial IoT. In *2022 20th Mediterranean Communication and Computer Networking Conference (MedComNet)* (pp. 84–93). IEEE.
15. Boateng, E.A., Bruce, J.W., Talbert, D.A.: Anomaly detection for a water treatment system based on one-class neural network. IEEE Access **10**, 115179–115191 (2022)
16. Mohammed, A.S., Anthi, E., Rana, O., Saxena, N., Burnap, P.: Detection and mitigation of field flooding attacks on oil and gas critical infrastructure communication. Comput. Secur. **124**, 103007 (2023)
17. Aydogan E, Yilmaz S, Sen S, Butun I, Forsström S, Gidlund M (2019, May) A central intrusion detection system for rpl-based industrial internet of things. In *2019 15th IEEE International Workshop on Factory Communication Systems (WFCS)* (pp. 1–5). IEEE.
18. Rahman, M.A., Hossain, M.S.: A deep learning assisted software defined security architecture for 6G wireless networks: IIoT perspective. IEEE Wirel. Commun. **29**(2), 52–59 (2022)
19. Kim, H.M., Lee, K.H.: IIoT malware detection using edge computing and deep learning for cybersecurity in smart factories. Appl. Sci. **12**(15), 7679 (2022)
20. Zhang Y, Yang C, Huang K, Li Y (2022) Intrusion detection of industrial internet-of-things based on reconstructed graph neural networks. *IEEE Transactions on Network Science and Engineering*.
21. Khan, I.A., Keshk, M., Pi, D., Khan, N., Hussain, Y., Soliman, H.: Enhancing IIoT networks protection: a robust security model for attack detection in internet industrial control systems. Ad Hoc Netw. **134**, 102930 (2022)
22. Khan, I.A., Moustafa, N., Pi, D., Sallam, K.M., Zomaya, A.Y., Li, B.: A new explainable deep learning framework for cyber threat discovery in industrial IoT networks. IEEE Internet Things J. **9**(13), 11604–11613 (2021)
23. Radoglou-Grammatikis P, Dalamagkas C, Lagkas T, Zafeiropoulou M, Atanasova M, Zlatev P, Sarigiannidis P (2022, December) False data injection attacks against low voltage distribution systems. In GLOBECOM 2022–2022 IEEE Global Communications Conference (pp. 1856–1861). IEEE.
24. Khan IA, Pi D, Abbas MZ, Zia U, Hussain Y, Soliman H (2022) Federated-SRUs: a federated simple recurrent units-based IDS for accurate detection of cyber attacks against IoT-augmented industrial control systems. IEEE Internet of Things Journal.
25. Louati, F., Ktata, F.B., Amous, I.: Big-IDS: a decentralized multi agent reinforcement learning approach for distributed intrusion detection in big data networks. Clust. Comput. (2024). https://doi.org/10.1007/s10586-024-04306-9

26. Nanjappan, M., Pradeep, K., Natesan, G., Samydurai, A., Premalatha, G.: DeepLG SecNet: utilizing deep LSTM and GRU with secure network for enhanced intrusion detection in IoT environments. Clust. Comput. (2024). https://doi.org/10.1007/s10586-023-04223-3
27. Chander, N., Upendra Kumar, M.: Enhanced pelican optimization algorithm with ensemble-based anomaly detection in industrial internet of things environment. Clust. Comput. (2024). https://doi.org/10.1007/s10586-024-04303-y
28. Alkhudaydi, O.A., Krichen, M., Alghamdi, A.D.: A deep learning methodology for predicting cybersecurity attacks on the internet of things. Information **14**(10), 550 (2023)
29. Radoglou-Grammatikis P, Zafeiropoulou M, Atanasova M, Zlatev P, Giannakidou S, Lagkas T, Sarigiannidis P (2023, June) False data injection attacks against high voltage transmission systems. In 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT) (pp. 324–329). IEEE.
30. Kelli V, Radoglou-Grammatikis P, Lagkas T, Markakis EK, Sarigiannidis P (2022, July) Risk analysis of DNP3 attacks. In *2022 IEEE International Conference on Cyber Security and Resilience (CSR)* (pp. 351–356). IEEE.
31. Sarker, I.H., Khan, A.I., Abushark, Y.B., Alsolami, F.: Internet of things (iot) security intelligence: a comprehensive overview, machine learning solutions and research directions. Mob. Netw. Appl. **28**(1), 296–312 (2023)
32. Amponis, G., Radoglou-Grammatikis, P., Lagkas, T., Mallouli, W., Cavalli, A., Klonidis, D., Sarigiannidis, P.: Threatening the 5G core via PFCP DoS attacks: the case of blocking UAV communications. J. Wireless. Com. Network. **2022**(1), 124 (2022)
33. Kelli V, Radoglou-Grammatikis P, Sesis A, Lagkas T, Fountoukidis E, Kafetzakis E, Sarigiannidis P (2022, May) Attacking and defending DNP3 ICS/SCADA systems. In 2022 18th International Conference on Distributed Computing in Sensor Systems (DCOSS) (pp. 183–190). IEEE.
34. Mladenov V, Chobanov V, Sarigiannidis P, Radoglou-Grammatikis PI, Hristov A, Zlatev P (2020, September) Defense against cyber-attacks on the hydro power plant connected in parallel with energy system. In 2020 12th Electrical Engineering Faculty Conference (BulEF) (pp. 1–6). IEEE.
35. Mohy-eddine, M., Guezzaz, A., Benkirane, S., Azrour, M.: Malicious detection model with artificial neural network in IoT-based smart farming security. Clust. Comput. (2024). https://doi.org/10.1007/s10586-024-04334-5
36. Sivasakthi, D.A., Sathiyaraj, A., Devendiran, R.: HybridRobustNet: enhancing detection of hybrid attacks in IoT networks through advanced learning approach. Clust. Comput. (2024). https://doi.org/10.1007/s10586-023-04248-8
37. Mallik, A.: Man-in-the-middle-attack: understanding in simple words. Cyberspace: Jurnal Pendidikan Teknologi Informasi **2**(2), 109–134 (2019)
38. Asad, M., Asim, M., Javed, T., Beg, M.O., Mujtaba, H., Abbas, S.: Deepdetect: detection of distributed denial of service attacks using deep learning. Comput. J. **63**(7), 983–994 (2020)
39. Polonijo B, Šuman S, Šimac I (2021, September) Propaganda detection using sentiment aware ensemble deep learning. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 199–204). IEEE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Gökçe Karacayılmaz is still continuing his PhD in Forensic Sciences at Hacettepe University. His current research interests include Cybersecurity about critical infrastructure, Operating Systems, Web technologies and its applications penetration testing and fuzzing.

Harun Artuner is still working as a lecturer in Computer Engineering at Hacettepe University. At the same time, he was the Founder of Hacettepe University Forensic Informatics Research and Application Center and is a member of the board of directors. He continues his studies on Digital Signal Processing, Speech Recognition, Embedded Systems, Artificial Understanding and Forensic Informatics.