IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 June 2023, accepted 28 June 2023, date of publication 3 July 2023, date of current version 24 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3291599

RESEARCH ARTICLE

APT Adversarial Defence Mechanism for Industrial IoT Enabled Cyber-Physical System

SAFDAR HUSSAIN JAVED[1], MAAZ BIN AHMAD[1], MUHAMMAD ASIF[2], WASEEM AKRAM[2], KHALID MAHMOOD[3], (Senior Member, IEEE), ASHOK KUMAR DAS[4], (Senior Member, IEEE), AND SACHIN SHETTY[5,6], (Senior Member, IEEE)

[1]College of Computing and Information Sciences, Karachi Institute of Economics and Technology (KIET), Karachi, Sindh 75190, Pakistan
[2]Department of Computer Science, Lahore Garrison University, Lahore 54810, Pakistan
[3]Graduate School of Intelligent Data Science, National Yunlin University of Science and Technology, Douliu 64002, Taiwan
[4]Center for Security, Theory and Algorithmic Research, International Institute of Information Technology at Hyderabad, Hyderabad 500032, India
[5]Department of Modeling, Simulation and Visualization Engineering, Virginia Modeling, Analysis and Simulation Center, Old Dominion University, Suffolk, VA 23435, USA
[6]Center for Cybersecurity Education and Research, Old Dominion University, Suffolk, VA 23435, USA

Corresponding authors: Ashok Kumar Das (iitkgp.akdas@gmail.com), Khalid Mahmood (khalidm.research@gmail.com), and Maaz Bin Ahmad (maaz@kiet.edu.pk)

ABSTRACT The objective of Advanced Persistent Threat (APT) attacks is to exploit Cyber-Physical Systems (CPSs) in combination with the Industrial Internet of Things (I-IoT) by using fast attack methods. Machine learning (ML) techniques have shown potential in identifying APT attacks in autonomous and malware detection systems. However, detecting hidden APT attacks in the I-IoT-enabled CPS domain and achieving real-time accuracy in detection present significant challenges for these techniques. To overcome these issues, a new approach is suggested that is based on the Graph Attention Network (GAN), a multi-dimensional algorithm that captures behavioral features along with the relevant information that other methods do not deliver. This approach utilizes masked self-attentional layers to address the limitations of prior Deep Learning (DL) methods that rely on convolutions. Two datasets, the DAPT2020 malware, and Edge I-IoT datasets are used to evaluate the approach, and it attains the highest detection accuracy of 96.97% and 95.97%, with prediction time of 20.56 seconds and 21.65 seconds, respectively. The GAN approach is compared to conventional ML algorithms, and simulation results demonstrate a significant performance improvement over these algorithms in the I-IoT-enabled CPS realm.

INDEX TERMS Advanced persistent threat, deep learning, cyber-physical systems, graph attention networks, graph neural networks, the Industrial Internet of Things.

I. INTRODUCTION

The Cyber-Physical Systems (CPSs) enabled by the Industrial Internet of Things (I-IoT) are software components that operate like hardware in automating industrial processes, collecting real-time data, and interacting with devices and sensors via Human-Computer Interfaces (HCI). This technology has several dimensions and aims to improve consistency, identify opportunities for progress, and exploit the untapped potential. By integrating the I-IoT sensors, data storage and integration, data analytics, and ML with CPSs, it is possible to enhance interoperability and coordination among various systems. The sensors gather data from different equipment and continually provide it to the system analytics. The ML algorithms then use this data to learn and refine the system's processes to reach optimal performance. The integration model of I-IoT with CPS is illustrated in Figure 1.

Furthermore, the fourth industrial revolution has brought about the concept of a ''smart factory'' through I-IoT, which enables cooperation among enterprise networks, supply chains, and manufacturing procedures [1]. Data is transmitted from machines to a top-level cloud server, providing a clear

The associate editor coordinating the review of this manuscript and approving it for publication was Razi Iqbal.

overview of the entire process. This system enables centralized control and remote monitoring of cyber assets located in harsh environments, connecting people, processes, and data in a more efficient, safer, and secure real-time information management system. Figure 2 illustrates the integration of I-IoT, IoT, and Industry 4.0. Moreover, there has been a significant expansion in the realm of I-IoT, which has led to a market size of more than USD 263 billion in 2021. It is predicted to surpass USD 350 billion by the year 2028 [2]. The reason behind this growth is the increased adoption of I-IoT platforms by large businesses that wish to keep up with the continuously developing technological landscape. As a result of this adoption, a huge volume of operational and transactional data is generated every second, which, when gathered and implemented in an I-IoT platform, can be transformed into real-time business insights and solutions. Figure 3 showcases the projected growth of the I-IoT market size from the year 2020 to 2028 [3].

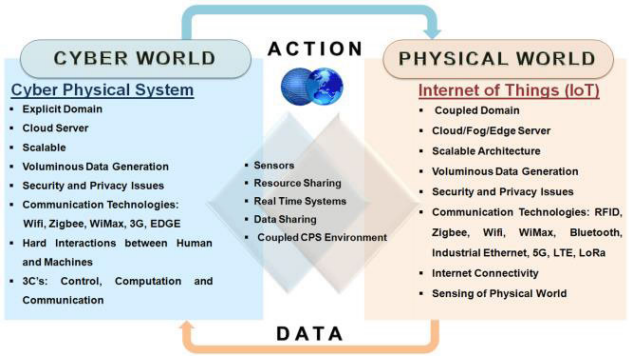

FIGURE 1. I-IoT and cyber-physical system integration model.

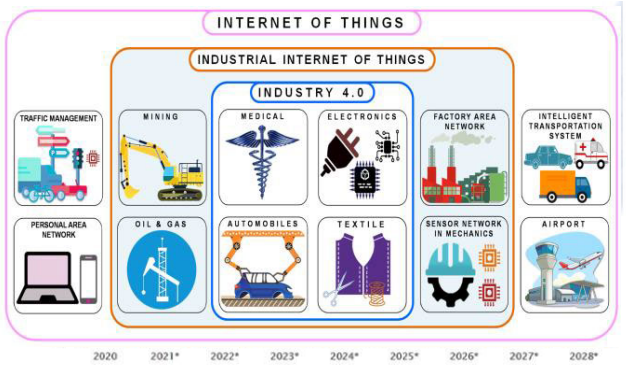

FIGURE 2. Integration of I-IoT and IoT with industry 4.0.

Although, the combination of I-IoT and CPS presents several potential benefits for society, however, this technology also faces numerous security challenges that must be addressed to ensure a reliable and scalable CPS environment. One of the most significant security challenges is the APT campaign, which uses multi-step attacks to pose severe threats to high-level information and hardware systems. These complex attacks make detection one of the major

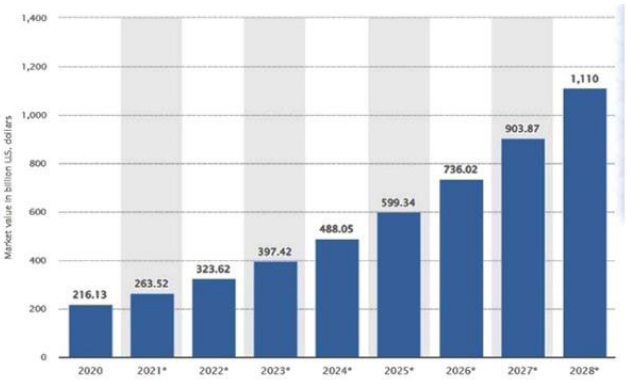

FIGURE 3. Projected market size of I-IoT 2020-28 [4].

security challenges facing industrial-scale CPS. As a result, these attacks can cause system interference, disrupt computer services, deny access to critical computer operations, and even lead to sabotage [4], [5]. Detecting APT attacks in I-IoT and CPS systems is difficult mainly because these systems are both scalable and heterogeneous. They also rely on isolated and complex data. Any exposure to these systems can lead to disastrous consequences, which is especially concerning because they are widely deployed systems. Conventional methods for understanding cyber threats typically rely on analyzing attack alerts to identify attack intentions. However, these methods are insufficient for dealing with the ever-changing threat landscape of complex cyber threats. It is essential to have a robust security system that can identify and prevent sophisticated APT attack campaigns quickly and accurately. This mechanism should be capable of countering APT and its advanced variants, reducing their destructive effects, and creating a cyber-situation comprehension system designed to detect APT attacks.

Numerous techniques based on ML and DL have been utilized to identify and categorize malware that exhibits intricate and harmful behaviors. The traditional security models, which rely on cryptographic techniques, are typically time-consuming and cannot efficiently handle large amounts of data, especially when dealing with sophisticated threats like APT campaigns. Consequently, these models are not commonly utilized in I-IoT environments, which generate huge amounts of data [6]. DL has demonstrated significant potential in various fields, including I-IoT. Its ability to handle intricate problems and its robustness have piqued the interest of numerous researchers, who have employed various DL-based algorithms in critical systems. These include but are not limited to Convolutional Neural Networks (CNN), Graph Neural Networks (GNN), Artificial Neural Networks (ANN), Boltzmann Machines, and Recurrent Neural Networks (RNN) to detect sophisticated threats [7].

In this research, a method of detecting and classifying APT attacks in the I-IoT-enabled CPS environment is proposed using Graph Attention Network (GAN). The proposed approach is designed to efficiently handle complex and

dynamic APT attacks. The method is evaluated on two publicly available datasets, the DAPT2020 malware, and the Edge I-IoT dataset, using standard performance metrics. The results show that the proposed GAN-based method achieves a high detection accuracy of 96.97% with a processing time of 20.56 seconds on the DAPT2020 malware dataset and 95.97% with a processing time of 21.65 seconds on the Edge I-IoT dataset. Furthermore, a comparison drawn with conventional ML techniques demonstrates the superiority of the proposed GAN-based method within the subject domain. The results of this study suggest that the proposed approach can provide effective protection against APT attacks in the I-IoT-enabled CPS environment.

The rest of the paper is structured as follows: Section II covers the APT and security challenges in I-IoT-enabled CPS. The related work in the field is presented in Section III. The methodology approach used in the study is outlined in Section IV. The results and analysis of the experiments are discussed in Chapter V, and the paper is concluded with suggestions for future research directions in Section VI.

II. APT AND SECURITY ISSUES IN I-IOT ENABLED CPS

The I-IoT plays a crucial role in the advancement of industry 4.0 by driving the smart manufacturing process. However, this progress has made the I-IoT a prime target for cyber attackers who face growing security challenges. Several well-known and unknown attacks have been observed on IoT devices with weak security and high vulnerabilities, including DDoS attacks, identity theft, MIME attacks, and network compromise [16]. However, due to limited resources, these devices are often unable to store enough information to defend against these attacks, making them vulnerable to attacks from botnets and APTs. These types of cyber-attacks can have disastrous effects on the entire organization which results in it becoming a prime target of cyber attackers, leading to significant security challenges. Many known and unknown types of attacks have been observed on IoT devices, which possess weak security and high vulnerabilities [8]. Due to various resource constraints, IoT nodes cannot often carry out efficient computer analysis with low information storage, leaving them vulnerable in terms of security. This provides opportunities for attackers to exploit the devices using known techniques, including but not limited to Distributed Denial of Service (DDoS) attacks, identity theft, MIME attacks, and compromised local networks [9]. In recent times, botnet and APT cyber-attacks have emerged as potential threats to I-IoT networks, which can have catastrophic effects on enterprise-wide organizations. The IoT devices are typically resource-constrained and are deployed across different locations, which makes it challenging to install computationally expensive attack detection systems on them. The traditional centralized cloud computing architecture is not capable of managing the high transmission overheads of attack detection systems. This approach could result in missing alerts and failing to detect certain attack intentions, particularly when it comes to complex APT attacks. APTs are a significant security challenge for I-IoT-enabled CPS due to their adaptable nature and refined exploitation methods.

A. APT THREAT CYCLE

The initial phase of an APT involves conducting reconnaissance activities to probe the targeted industrial network for vulnerable and exploitable components. This is done by gathering information about the network from general internet searches or using social engineering tools for propagation. In the weaponization phase, the attacker creates a malicious document paired with a customized phishing email or employs a new strain of self-replicating malware. These are then distributed through various means such as malware laden email, Wi-Fi, or other entry points into the core network. The delivery phase focuses on transmitting the malware to the intended industrial network, taking advantage of weak account and password management measures as a gateway for intrusion. Once inside the network, the exploitation phase begins, with the attacker using the infiltrated malware to exploit vulnerabilities within the target network. In the installation phase, the APT malware is installed on the processing layer machines, establishing a foothold. The Command and Control (C&C) phase serves as a post-compromise layer, enabling the attacker to control the compromised systems through an external C&C system. The communication between the compromised host and the C&C system is typically encrypted to avoid detection. The final phase, known as actions on objectives or data exfiltration, involves the cyber attacker gaining access to the organization and executing actions to achieve their objectives, often involving the theft or exfiltration of high-value data [10]. Within the context of the IoT enabled CPS domain, APT attacks the core CPS to exploit various areas of vulnerability. These may involve altering the sensor and actuator thresholds, manipulating the network connections between controllers and actuators, and disrupting the network connectivity between sensors and the Human Machine Interface (HMI) connected controllers [11]. Existing state-of-the-art system defenders are unable to withstand the highly sophisticated, covert, and deceptive APT attacks. They are developed by cyber adversaries having technical expertise that use complicated attack methods and take advantage of various intrusion programs to achieve their attack objectives. Often referred to as ''one-day exploits'' the attacker continue to have other attack objectives even if the critical system is breached [12]. Figure 4 depicts the most significant security obstacles faced by Cyber-Physical Systems enabled by I-IoT.

In summary, I-IoT enabled CPS in industrial environments are critical for mission-critical operations, but they have limitations when it comes to security. Therefore, it is crucial to develop a real-time security monitoring mechanism that specifically focuses on preventing unauthorized and malicious users from accessing industrial critical systems. Traditional security models rely on cryptographic techniques, which require significant processing time to analyze large

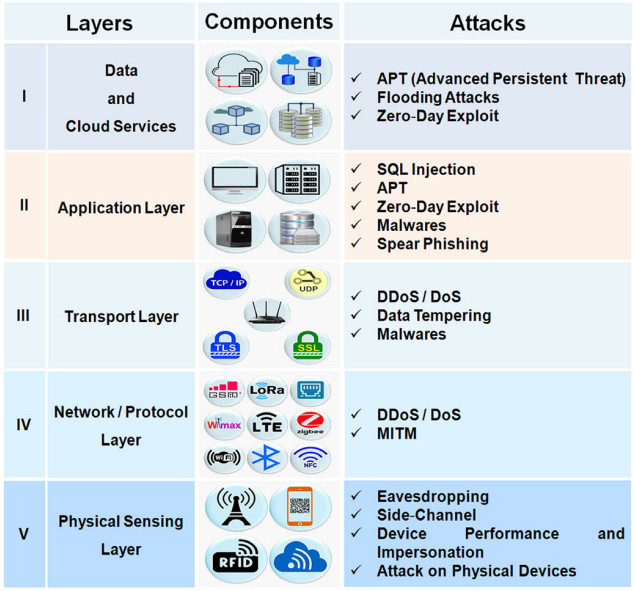

FIGURE 4. Security challenges in I-IoT enabled CPS.

volumes of data and detect complex threats like APTs. As a result, these models are not widely adopted in I-IoT environments due to the extensive size of data involved. In order to minimize potential damage, it is necessary to promptly identify security threats and take appropriate actions. By promptly identifying and analyzing APT data traffic within the I-IoT domain, organizations that deploy I-IoT sensors can effectively protect their most valuable assets and data from digital disruption in industrial processing units.

III. RELATED WORK

Identifying and accurately categorizing APT signatures in the context of the I-IoT-enabled CPS domain is a difficult challenge that many researchers and solution providers have attempted to tackle. As a result, several efficient intrusion detection systems have been proposed over time. A variety of signature-based and behavior-based security frameworks have been put forward to detect and classify cyber threats, including APT which are discussed.

Azizjon et al. [13] propose a novel technique for identifying malware that employs edge computing and DL, more specifically, the CNN model. The CNN is utilized to translate the binary file of the malware into images composed of pixels, and this methodology has demonstrated an accuracy rate of 98.93% on the Malimg dataset intended for the I-IoT environment. The system involves the distribution of considerable amounts of traffic data generated by smart factories' I-IoT to edge servers for processing by DL. The system consists of three layers, namely the edge device layer, edge layer, and cloud layer. The edge-based DL approach comprises four functions, which include model training and testing, model deployment, model inference, and transmission of training data. While the system has demonstrated excellent accuracy on publicly available datasets, it must be tested in real-time situations on APT datasets to ensure its effectiveness.

In their research, Huang and Zhu [14] have presented a game-theoretic methodology for developing proactive and cross-layer defenses against APT in a CPS environment. This approach involves each player creating a belief about unknown variables and utilizing Bayesian updates to learn private information and minimize uncertainty. By analyzing the Perfect Bayesian Nash Equilibrium (PBNE), the authors have provided the defender with an effective countermeasure against strategic attacks at multiple stages. They have also introduced a nested algorithm that alternates between forward belief updates and backward policy computation, rapidly converging to the E-PBNE and providing a consistent set of beliefs and policies for identifying complex malware attacks. The experimental results have demonstrated that a sophisticated defense can receive a 56% higher payoff compared to a primitive defense. However, it is necessary to validate the efficacy of this approach on real-time APT datasets to demonstrate its effectiveness in I-IoT-enabled CPS.

Tamy et al. [15] present ML model for classifying and predicting cyberattacks by using different conventional algorithms, including Naive Bayes, SVM, J48, and Random Decision Forest (RDF), on a ''10% Random Sample Gas Pipeline'' dataset. The objective of the research is to identify the optimal algorithm for detecting and predicting cyberattacks in the CPS so that appropriate preventive actions can be taken to reduce the risk of intrusion. The findings of the study showed that RDF is the most effective, with a remarkable accuracy rate of 99.30%. However, this method should still be tested using the APT dataset in a real-time I-IoT CPS environment.

Qian et al. [16] introduced an IDS-based system designed to detect Man-in-the-Middle (MITM) and Replay attacks in cyber and physical systems that utilize the Modbus TCP Protocol. The system uses a validation process to identify malicious activity and prevent harm to the physical system caused by MITM, Replay, and Zero-day attacks. A non-parallel hyper-plane fuzzy classifier utilizing SVM is used to detect DoS (SYN flood) attacks in the cyber domain. The system employs a 41-dimensional dataset containing 2200 samples gathered from a Supervisory Control and Data Acquisitions (SCADA) system combined with Modbus and TCP protocol traffic data. Although this system is capable of detecting the aforementioned attacks, it cannot determine the location or type of the attacks, making it unsuitable for detecting APT-type attacks.

Gao et al. [17] developed an IDS that employs RNN and Long-Short Term Memory (LSTM) with MTM and MTO architectures in the CPS domain network. The system utilizes two different datasets, one consisting of correlated data and the other uncorrelated data, collected in real-time, and extracts relevant features. The results showed that the MTO architecture is effective in detecting sequentially uncorrelated attacks, achieving a 90% F1 score accuracy, and it performed even better in detecting temporally coordinated attacks.

Stewart et al. [18] conduct a study to examine the impact of changes in the network architecture of the CPS system on the performance of an IDS using a one-class Support Vector Machine (OCSVM). The system is designed to be adaptive and can adjust to real-time situations. The effectiveness of the system is evaluated using traces from a hybrid ICS testbed and the NSL-KDD dataset.

George and Thampi [19] addressed the security vulnerability concerns in the I-IoT network and its devices and presented a graph model for its representation. The model serves as a security framework for evaluating and reducing the risk of network traffic. It provides a multi-faceted and multi-host attack detection system that focuses on the chain of vulnerabilities in I-IoT networks. The approach extracts security-related parameters from the graph-based model by eliminating high-risk attack paths and attack paths with low hop lengths and hot spots.

Teixeira et al. [20] develop a model that employs five traditional ML algorithms for detecting cyber-attacks on both online and offline systems. The model consists of RDF, DT, LR, NB, and KNN. The ML models that have been trained are subsequently implemented in the network by utilizing real-time network traffic. The performance of the model during the training and testing phase was compared with the results obtained from its real-time deployment online. The results showed that RDF achieved 100% and 99.89% accuracy with a False Alarm Rate (FAR) of 0.00 on online and offline systems, respectively.

Lin and Nadjm-Tehrani [21] introduce a method for modeling the timing characteristics of spontaneous events in the IEC-60870-5-104 network and using the model for detecting anomalies in the CPS domain. The method is tested using a real-time power utility dataset that introduced timing effects to detect two types of attacks. One attack causes time-based anomalies that result in the malfunction of edge devices, while the other involves sporadic, stealthy attacks. The results are promising, with 99.99% of all persistence attacks detected. The approach needs to be tested using APT datasets for its efficacy in a CPS-I-IoT environment.

Zhou et al. [22] present a framework for behavior-based anomaly detection, which collects information to create three different normal behavior baselines from various dimensions. The framework employs a transparent network snooping mechanism on the ICS system components. Passive recognition methods use PCAP files generated by the Wireshark tool and online sniffers for data analysis. However, this approach has a drawback in that the data packet tampering and logical attacks go undetected, and the framework needs to be tested for APT detection in I-IoT and CPS domains.

Hnamte and Hussain [32] propose a framework called MLAPT for detecting and predicting APT signature attacks. This framework is divided into three phases: threat detection, alert correlation, and attack prediction. The first phase aims to reduce false positive rates, the second phase uses a machine learning-based prediction module to monitor network repository data, and the third phase quickly captures attacks. The framework achieved an accuracy of 84.8% on a generic system.

Veličković et al. [33] put forth a system for detecting APT that categorizes Command and Control (C & C) communications. The classification system showed impressive results when tested on a publicly available dataset, with a True Positive Rate (TPR) of 83.3%. However, this approach is susceptible to evasion if infected hosts connect to a C & C domain. Additionally, the whole APT life cycle may not be detected if the signatures of the C & C remain unnoticed.

You et al. [34] propose a method for detecting APT using the spear-phishing technique. A mathematical computation filter is used to identify spam emails using tokens of detection algorithms, separating legitimate and spam emails. Nonetheless, this approach is limited because it only employs a single step for identifying APT traffic, which renders it unsuitable for deployment in the I-IoT realm.

Yu et al. [23] propose a DL-based proactive APT detection scheme in I-IoT that uses bidirectional encoder representations from transformers (BERT) scheme that detects APT attack sequences. The APT attack sequence is optimized to ensure the model's long-term sequence judgment effectiveness. The approach is authenticated on a dataset gathered through various equipment manufacturers and categorized into five simulated attack categories: ''NORMAL'', ''PROBE'', ''DOS'', ''U2R'' and ''R2L''. The scheme provides an accuracy of 99%. The scheme needs to be validated computationally for the I-IoT environment.

Siniosoglou et al. [24] present a unified DL-based anomaly detection and classification approach that targets APT and security threats in I-IoT enabled smart grid environment. The proposed IDS called MENSA (Anomaly Detection and classification) adopts a novel Auto-Encoder-Generative Adversarial Network (A-EGAN) architecture for detecting operational anomalies and classifying Modbus/TCP and DNP3 including APT cyberattacks. The scheme is validated on various datasets that include Modbus/TCP network flows, DNP3 network flows and operational time-series electricity measurements data and provides accuracy, TRP, and FPR of 0.947%, 0.812, and 0.036 respectively. Although the approach provides convincing results, however, validation for APT data traffic needs to be ascertained including training and testing time of the framework.

Kumar and Thing [25] have introduced RAPTOR, an APT detection system that has been specifically developed for I-IoT environments. RAPTOR identifies and connects attack stages obtained from an APT Attack Invariant State Machine using optimal data sources that are selected for each stage. The correlated attack stages are used to produce a concise, high-level APT Campaign Graph that can track the progress of the APT campaign and implement suitable mitigation measures. Performance evaluations of RAPTOR show that it can detect APT campaigns, modeled after real-world attacks, with high precision and low false positive and negative

TABLE 1. Summary of related work.

Authors	Objectives and Methods	Datasets	Observations
Kim et al. [13]	A deep learning-based malware detection system that utilizes CNN has been implemented for use in the I-IoT environment. This system transforms malware binary files into image pixels for detection at the edge of the network.	Malimg dataset	The approach has a high accuracy of 98.93% on a publicly available dataset, but it needs to be tested in a real-time environment using an APT dataset.
Huang et al. [14]	Implements a game theory-based framework to comprehensively address APT in cyber-physical systems through proactive, cross-layer defenses. Uses an iterative algorithm to calculate the ideal Bayesian Nash Equilibrium and tests it using the Tennessee Eastman process as a benchmark.	-	The approach results in a 56% increase in payoff compared to a basic defense mechanism. However, its effectiveness in an I-IoT-enabled CPS environment still needs to be confirmed through testing on a real-time APT dataset.
Tamy et al. [15]	Employs the Naive Bayes, Support Vector Machine (SVM), J48, and Random Decision Forest (RDF) algorithms for categorizing data.	10% Random Sample Gas Pipeline dataset	The Random Decision Forest shows better results than other machine learning algorithms with an accuracy of 99.30%. The effectiveness of this approach in the I-IoT domain still needs to be tested.
Qian et al. [16]	A system that identifies MITM and replay attacks in both the cyber and physical aspects of CPS is proposed. Different data classification methods, such as NHFC, SVM, OCSVM, FCMSVM, GENFIS, and NHFC, are applied in the system.	A dataset with 41 different variables and 2200 samples was gathered from a CPS system using the Modbus/TCP protocol and network traffic data.	The IDS simulation yields an accuracy of 97.2%, however, it lacks visibility in terms of the location and type of attacks, making it inappropriate for detecting APT attack campaigns.
Gao et al. [17]	A DL-based IDS is proposed that uses RNN and LSTM architectures with MTM and MTO connections in a SCADA network.	Employs two distinct datasets, one of which is correlated and the other uncorrelated, that were gathered in real-time.	The approach detects temporally correlated attacks with an accuracy of 90% as measured by the F1 score. However, its effectiveness in the I-IoT domain still needs to be validated.
Stewart et al. [18]	The researchers introduced an IDS that employs the OCSVM technique and can adapt to real-time scenarios. To test its efficiency, they evaluated the adaptive ID using trace data gathered from a hybrid Industrial Control System (ICS) testbed.	NSL-KDD dataset	The effectiveness of the adaptive framework must be assessed in real-world scenarios in the I-IoT domain by testing it with actual APT data.
George et al. [19]	An IDS that uses Graph Neural Network is proposed. The method involves extracting security parameters by employing a graph-based model to eliminate vulnerable attack paths.	The approach makes use of Common Vulnerability Scoring System (CVSS) data, which is an open security standard for identifying specific characteristics and consequences of vulnerabilities.	The approach requires validation in real-world conditions in the I-IoT domain.
Teixeira et al.[20]	Applies conventional ML algorithms such as Random Forest, Decision Tree, Logistic Regression, Naive Bayes, and KNN to classify anomalous data.	Dataset gathered through live network stream.	The effectiveness of this approach needs to be assessed by comparing it with deep learning-based methods in detecting APT attacks in I-IoT-enabled cyber-physical systems using a dataset of such attacks.
Lin et al. [21]	The approach involves modeling the timing of unexpected events in an IEC-60870-5-104 network and using the model to detect anomalies in the field of CPS.	The approach has been evaluated using actual data from a power utility system.	The approach has an accuracy rate of 99.99% on an IoT dataset, but it needs to be evaluated for its ability to detect persistent attacks in I-IoT-enabled CPS by testing it on an APT dataset.
Zhou et al. [22]	The framework for the behavior-based anomaly detection mechanism collects information to create three different normal behavior patterns from multiple perspectives to detect an attack.	The approach uses PCAP data files generated by Wireshark and direct online sniffers for data analysis.	The proposed framework approach has limitations in detecting fake and logical attacks, as well as tampering with data packets. It needs to be tested for its ability to detect APT attacks in the IoT domain.

TABLE 1. *(Continued.)* **Summary of related work.**

Ghafir et al. [29]	A framework named MLAPT has been presented for detecting and predicting APT. It consists of three components: threat detection, alert correlation, and attack prediction in the network.	Customized ML Dataset	The accuracy of the method is 84.8%. However, further testing is required to determine its effectiveness in detecting APTs in the domains of I-IoT and CPS using the APT dataset.
Wang et al. [30]	A system that detects the Command and Control (C&C) stage of APTs by analyzing the signature of the C&C adversary.	The APT dataset is available to the public and is provided by the Los Almos Laboratory.	The approach attains a true positive rate of 83.3%. However, if the infected hosts are connected to an external network, it could result in a single point of failure, and the system could be evaded.
Chandra et al. [31]	The approach utilizes spear-phishing and mathematical computation filter techniques to eliminate unwanted spam emails.	Spam Database	The approach is not suitable for the IoT domain as it has the potential to result in a single point of failure.
Yu et al. [23]	A DL-based proactive APT detection system for the IoT has been developed. It uses the Bidirectional Encoder Representations from Transformers (BERT) scheme to identify APT attack sequences.	A dataset was collected from multiple equipment manufacturers and it was divided into five categories of simulated attacks: "NORMAL", "PROBE", "DOS", "U2R", and "R2L".	The approach has an accuracy of 99%, however, it still needs to be computationally validated in the context of the I-IoT environment.
Siniosoglou et al [24]	The MENSA-IDS is suggested, which applies an innovative Auto-Encoder Generative Adversarial Network (GAN) structure for identifying uncommon events and classifying Modbus/TCP and DNP3 cyberattacks, such as APTs.	Data on Modbus/TCP network flows, DNP3 network flows, and time-series measurements of electricity operations.	The system achieves an accuracy of 0.947%, TPR of 0.812, and FPR of 0.036. However, it still needs to be validated for APT data traffic.
Kumar et al [25]	RAPTOR is a system that identifies and connects the various stages of an APT attack (tailored to the I-IoT) by utilizing multiple data sources. It creates a comprehensive APT attack graph that can be utilized by security experts for analyzing and countering the attack.	The synthetic APT campaign was incorporated into the CSE-IDC2018 intrusion detection dataset.	The system boasts a precision rate of 0.996. However, one of its limitations is that APT malware may try to avoid detection by RAPTOR by slowing down the scanning process (e.g. during the Discovery or Fieldbus scanning stages) or altering the interval between scans to confuse the trained machine learning algorithm.
Xiao et al [27]	Sybil-based collusion attacks (SCA) in the context of IIoT-FL systems are proposed for detecting malicious attacks. The malicious participants use label-flipping attacks to complete local poisoning training and virtualize multiple Sybil nodes to make the local poisoning models aggregated with the greatest possibility during aggregation. Overall, the study highlights the importance of detecting SCA in IIoT-FL systems to ensure the security and reliability of these systems.	-	The system needs to be tested on APT datasets for its effectiveness and real-time accuracy.
Dhelim et al [28]	An Autonomous Trust Management System (TMS) called Trust2Vec is proposed which is capable of mitigating large-scale trust attacks that involve hundreds of malicious devices.	Simulated Dataset	The proposed system achieves a mitigation rate of up to 94%. However, its effectiveness needs to be validated in the APT attack dataset.
Hnamte et al[32]	An intelligent and efficient NIDS based on CNN and LSTM model is proposed, for the detection of attacks. The model is trained on real-time traffic datasets and achieves high-end accuracy on network traffic.	CICIDS2018 and Edge_IIoT datasets.	The approach achieves very high accuracy of 100% and 99.64% on relevant datasets, however, its efficacy in terms of performance needs to be validated for the I-IoT domain.

rates in the I-IoT domain. The authors have used synthetic APT campaigns inserted into the CSE-IDC2018 intrusion detection dataset to evaluate the approach, which yielded a precision accuracy of 0.996%. However, further validation of this approach in a real-time I-IoT environment is necessary.

Yao et al. [26] presented a hybrid intrusion detection system for Edge-Based I-IoT relying on ML-aided detection for edge-based I-IoT, innovations in detection methods, and system architecture. At the detection method level, the lightweight LightGBM DL algorithm using CNN algorithms

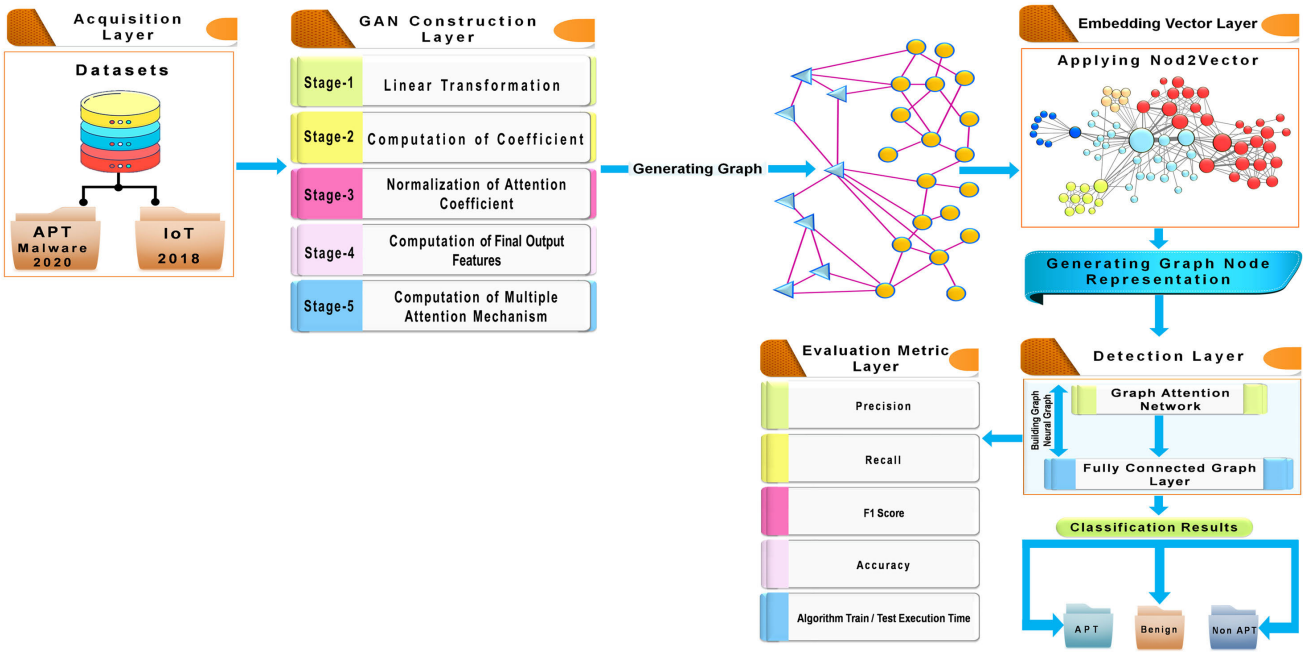

FIGURE 5. The proposed adversarial system.

is utilized in the lower level and upper level of the network, respectively. Xiao et al. [27] suggest Sybil-based collusion attacks (SCA) in the I-IoT-FL system for the detection of malicious attacks.

Dhelim et al. [28] suggested a Trust Management System (TMS) called Trust2Vec which is an integral component of any IoT network. Trust2Vec is capable of mitigating large-scale trust attacks that involve hundreds of malicious devices. The system uses a random-walk network exploration algorithm that navigates the trust relationship among devices and computes trust network embedding. To detect large-scale attacks such as self-promoting and bad-mouthing, the system proposes a network embedding community detection algorithm that identifies and blocks communities of malicious nodes. The proposed system achieves a mitigation rate of up to 94% in various network settings.

Hnamte et al. [57] propose an intelligent and efficient Network Intrusion Detection System (NIDS) based on DL. The focus is on developing a DL-based IDS for the detection of attacks. To train the model, real-time traffic datasets, specifically CICIDS2018 and Edge_IIoT datasets, are utilized. The performance of the model is evaluated using multiclass classification, and it achieves an accuracy rate of 100% when trained and tested with the CICIDS 2018 dataset.

Similarly, when trained and tested with the Edge_IIoT dataset, the model achieves an accuracy rate of 99.64%. These results demonstrate the effectiveness of the proposed DL-based IDS in detecting attacks in network traffic. Although, the approach achieves high accuracy on two diversified datasets, one related to the I-IoT domain, however, the approach still needs to be validated for the I-IoT domain. In short, current methods focus on identifying anomalous traffic and APT attacks in I-IoT-enabled CPS using various DL and autonomous techniques. However, their effectiveness in detecting modern cyber threats like APT in in the I-IoT-enabled CPS, domain is questionable. Table 1 summarizes the related work.

Conclusively speaking, these techniques have been tested in generic and SCADA systems using generic and specific datasets, but they cannot provide real-time recognition of genuine APT threats, limiting the security administrator's decision-making ability. Moreover, detecting all APT phases and balancing between false positive and negative rates also remains a challenging task. Furthermore, the suitability of these models in detecting APT malware traffic needs to be tested in mission-critical cyber complex domains where computational resources are limited. Moreover, they have not been tested on real-time systems with APT-specific datasets, as any false positive or false negative can have a significant impact on the complex system. Therefore, there is a need for a computationally efficient APT malware detection system that can quickly identify APT attacks and protect the mission-critical infrastructure of I-IoT-enabled CPS. To achieve this goal, this study presents a computationally effective DL-based APT malware detection and classification system that uses a GAN model on APT and IoT-specific datasets suitable for complex I-IoT-enabled CPS. The proposed GAN-based model combines the strengths of Neural Networks (NN) and node feature generators on DAPT2020 malware and Edge I-IoT datasets to cover all phases of the APT cycle in the cyber complex domain.

IV. METHODOLOGY

Recently, the Graph Attention Network (GAN) has garnered substantial interest among cyber security researchers due to its unique capabilities. Unlike other NN models, such as

CNNs, Recurrent Neural Nets (RNNs), or Auto-encoders, the GAN allows each node to be associated with a label that enables the prediction of unknown nodes by leveraging the information contained in the edges connecting neighboring nodes. This is an advantage that the other NN models cannot replicate due to their inability to model graph structures [33]. While traditional NN models such as CNNs, RNNs, or Auto-encoders are effective at identifying patterns in data such as text, images, or video, they are not well suited for handling graph structures, which are interconnected by various nodes. To address this limitation, GAN is developed as a DL method specifically designed for handling graph data. GANs are applied directly to graph structures, offering the advantage of making predictions at the node, edge, and graph levels. This makes them a suitable solution for analyzing graph-structured data [34]. Additionally, GAN can also be utilized for training networks with less computational expense, using methods like Spatial Graph Convolution Networks (SGCN) and Spectral Graph Convolution Networks [35]. In our proposed approach, GAN is employed to classify and detect intricate APT malware in graphical data. It works on the principle that each adjacent node contributes equally to producing the central node representation. The attention mechanism of GAN allocates varying importance to each neighboring node's contribution, rendering it more reliable and effective than other NN [36]. Moreover, to prevent overfitting during model training, the structure incorporates dropout and regularization layers after the convolution and global pooling layers. Node2Vec embedding is also integrated as a node feature, which utilizes parameters to determine the pace at which adjacent nodes are encountered in graph traversals. Figure 5 depicts the proposed structure of our system in the realm of IoT-enabled CPS.

The system is composed of five layers: the data acquisition layer, the GAN construction layer, the embedding vector layer, the detection layer, and the model evaluation layer. The system is built in four stages, starting with convolution layers that create generalized feature representations of a specific size for each node. To decrease the dimensions, a global pooling layer is employed. Dropout and regularization stages have been included after the convolution and global pooling layers to prevent over-fitting while training the model. The Adam optimizer is preferred over other optimizers such as AdaBelief, Adagrad, and Rmsprop due to its faster convergence and higher accuracy [37]. The output is generated in the global pooling layer and presented as a softmax layer to determine whether the network traffic is APT malware, benign, or normal data traffic. The GAN model's hyper-parameters are listed in Table 2.

TABLE 2. Hyper-parameters of the proposed graph attention network model.

Attributes	Numeric
#epoch	100
#Hidden Neurons	16,32,64,128
# Levels	3
Dropout coefficient	0.2
L2-regularization coefficient	0.01
Optimization algorithm	Adam

A. DATASETS

As complex cyber threats are becoming increasingly intricate in their tactics, techniques, and procedures (TTP), many modern cyber-attacks can be characterized by their TTP that concentrates on analyzing and profiling distinct threat vectors.

In the same manner, most of today's widespread malware threats follow similar actions to those of APT attack campaigns [38]. To evaluate our proposed system, we utilized the following publicly accessible datasets: -

B. DAPT2020[1]

Complex cyber threats are becoming more multifaceted in their tactics, techniques, and procedures (TTP). The DAPT2020 dataset [39] facilitates and enables a better understanding of the relationship between the APT groups and TTP. In this dataset, the analysis of attack behavior at the interface and inside the network level incorporates four major APT attack phases that include Intelligence gathering, Penetration, Network propagation, and Data outflow. Additionally, the network flow characteristics in DAPT 2020encode numerous latent factors, which include versatility and stealthiness, which are core features of APTs. The dataset captures the various aspects of real-world APT attacks, which include attack behavior both at the interface and inside the network. The threat model used for the creation of the APT dataset incorporates the four main phases of an APT attack reconnaissance, foothold establishment, lateral movement, and data exfiltration. The traffic features in the dataset encode several latent characteristics, such as adaptability and stealthiness, of APTs spanning all the stages of an APT. Generalized APT attack phases mapped to its real-time activity are summarized in Table 3.

C. EDGE–IIOT[2]

We utilized a comprehensive dataset called Edge-IIoT [40], [41], which is focused on cyber security in IoT and I-IoT applications. The IDS systems based on ML algorithms were developed using a dataset called Edge-IIoT. This dataset includes data from various IoT devices such as digital sensors for temperature and humidity sensing, ultrasonic sensors, water level detection sensors, pH sensor meters, soil moisture sensors, heart rate sensors, and flame sensors. It contains 14 attacks that target IoT and I-IoT connectivity protocols and are divided into five categories: DoS/DDoS attacks, information gathering, man-in-the-middle attacks, injection attacks, and malware attacks. The dataset also

[1]https://www.kaggle.com/datasets/sowmyamyneni/dapt2020
[2]https://ieee-dataport.org/documents/edge-iiotset-new-comprehensive-realistic- cyber-security-dataset-iot-and-iiot-applications

TABLE 3. Network activities with mapped APT attack phases on DAPT2020 dataset.

Generalized APT Attack Phases	Activity	Tools Used	Details
Normal Traffic	Network Scan	ping, dig, GET, POST, curl, browsing, files upload, download	Establishing a standard level of regular traffic according to users' actions
	Application Scan		
	Account Brute Force		
Reconnaissance	Network Scan	nmap, webscarab, sqlmap, dirbuster, nikto, burpsuite, application account discovery tools	Scouting on the public network to identify vulnerabilities, the structure of directories, Poor authentication, and authorization protocols
	Application Scan		
	Account Brute Force		
Foothold Establishment	CSRF	PHP reverse shell, netcat, SQL vulnerability exploitation, authentication bypass, metasploitable	An attack was carried out on the I-IoT network, which involved using a PHP reverse shell via the DVWA, uploading a file, and adding malicious users.
	SQL Injection		
	Malware Download		
	Backdoor		
	Reverse Shell		
	Command Injection		
Lateral Movement	Internal Scanning	A scan of the local network using Nmap to identify the vulnerability in vsftpd 2.3.4, weak ssh authentication, a mysql script exploiting CVE-2012-2122, and the metasploit framework	The exploration of the internal network from the hacked virtual machines (public VM) and gaining access to critical local systems.
	Account Discovery		
	Password Dumping		
	Credential Theft		
	Creation of user accounts		
	Privilege Escalation		
Data Exfiltration	Data Theft	Exfiltration of data to the C&C server, exploiting SMB vulnerability (CVE-2017-7494) to attain higher privileges, through Google Drive, PyExfil, FTP, and SCP	The transfer of files from a local machine to a remote server using the FTP put method, wput to a remote location using an anonymous user, scp to transfer large files to the remote server, and web-based uploads to Google Drive.

includes features extracted from various sources such as system resources, logs, alerts, and network traffic. The dataset includes 2,219,201 attacks, with 1,615,643 considered regular and 603,558 classified as attacks. Tables 4 and 5 provide information on the attack scenarios in the Edge-IIoT dataset and the selected categories, respectively.

D. ALGORITHMIC WORKFLOW OF GAN MODEL

Before feeding the data into the NN classifier, it is pre-processed and normalized. Careful data analysis is crucial for accurately predicting I-IoT traffic as APT malware, benign APT, or normal data traffic. So, the first step is to organize the data in a way that is compatible with input for DL classifiers. During the data de-noising process, any datasets containing missing, infinite, or NAN values including unexpected values are removed. Types of features such as numerical and categorical data are also identified, and categorical data are converted to numerical data through the process of label encoding.

The layer takes a set of node features, denoted as h = {h1, h2, …, h N}, where each hi belongs to RF, representing the features of each node. N represents the total number of

TABLE 4. Attack scenario – edge IIoT dataset.

Attack Classification	Category	Security Weakness	APT Attack Phases
Dos / DDoS Attack	TCP SYN Flood DDOS Attack	Make target edge IoT sever unavailable to legitimate requests	Foothold Establishment
	UDP flood DDoS attack	Overcome the dispensation and response abilities of target devices	
	HTTP flood DDoS attack	Exploits impersonated – legitimate HTTP GET and HTTP Post requests to IoT application	
	ICMP flood DDoS attack	The IoT Edge servers become inaccessible to normalized network data flow by flooding through ICMP echo requests.	
Information Gathering	Port Scanning	Discovers open doors and weak security links in edge-based IoT network	Reconnaissance
	OS Fingerprinting	Analyzing IoT network data flow to identify the shortfalls of targeted IoT devices	
	Vulnerability scanning attack	Identify IoT system vulnerabilities	
MIME Attack	DNS Spoofing attack	The capture of communication between targeted IoT devices and the Domain server	Data Exfiltration
	ARP Spoofing attack	Linking the attacker's MAC address with the IP address of an IoT device / Edge server	
Injection Attack	XSS Attack	Transmitting a harmful script to an unsuspecting user that can gain access to sensitive information, session tokens, cookies, and other resources	Lateral Movement
	SQL injection	Insertion of malicious SQL query to read, insert, update, or delete sensitive data from an IoT database	
	Uploading attack	Uploading files that contain malware command and control data	
Malware Attack	Backdoor attack	Creating backdoors to gain control over vulnerable components of an IoT network	
	Password Cracking attack	Identifying a password that has been forgotten for an IoT device to gain unauthorized access to IoT resources	
	Ransomware attack	Manage access to IoT data or systems by utilizing encryption	Infiltration

nodes, and F represents the number of features in each node. The layer generates a new set of node features, also denoted as h = {h1, h2, …, h N}, where each hi belongs to RF, as the output. The objective is to have sufficient expressive power to transform the input features into higher-level features, which requires at least one learnable linear transformation. As a first step, a shared linear transformation is applied to each node using a weight matrix, W, which has dimensions F×F. Subsequently, the layer employs a self-attention mechanism on the nodes. This mechanism, represented by a function a: RF×RF → R, computes attention coefficients by Eq 1:

$$e_{ij} = a\left(W\vec{\boldsymbol{h}_i}, W\vec{\boldsymbol{h}_j}\right) \tag{1}$$

The importance of the node j's features to node i is determined by the attention coefficients. In the general formulation of the model, every node can attend to all other nodes, without considering the structural information of the graph. However, we incorporate the graph structure by applying masked attention. This means that we only compute the attention

TABLE 5. Selected classes – edge IIoT dataset.

S/No.	Classes	S/No.	Classes
1.	Normal	8.	Password attack
2.	Backdoor attack	9.	Port_Scanning
3.	DDoS_HTTP attack	10.	Ransomware attack
4.	DDoS_ICMP attack	11.	SQL injection attack
5.	DDoS_TCP attack	12.	Uploading attack
6.	DDoS_UDP attack	13.	Vulnerability_scanner attack
7.	Fingerprinting attack		

coefficient e_{ij} for nodes j that belong to the neighborhood of node i, denoted as Ni. In our experiments, the neighborhood Ni specifically includes the first-order neighbors of node i, including node i itself. To ensure that the attention coefficients can be compared across different nodes, we normalize them using the softmax function. This normalization is performed across all possible choices of node j within the neighborhood N_i computed by Eq 2:

$$\alpha_{ij} = softmax_j\left(e_{ij}\right) = \frac{\exp\left(e_{ij}\right)}{\sum_{k\epsilon N_i}\exp\left(e_{ik}\right)} \tag{2}$$

In our experiments, the attention mechanism, denoted as a, is implemented as a single-layer feed-forward neural network. It is parametrized by a weight vector belonging to the real number R and has a dimension of 2F. The non-linearity used in this mechanism is the Leaky ReLU, with a negative input slope of $\alpha = 0.2$. Expanding the attention mechanism further, the computed coefficients can be expressed as follows in Eq 3:

$$\alpha_{ij} = \frac{\exp\left(LeakyReLU\left(\boldsymbol{a}^{\rightarrow T}[W\boldsymbol{h}_i^{\rightarrow}||W\boldsymbol{h}_j^{\rightarrow}]\right)\right)}{\sum_{k\epsilon N_i}\exp\left(LeakyReLU\left(\boldsymbol{a}^{\rightarrow T}[W\boldsymbol{h}_i^{\rightarrow}||W\boldsymbol{h}_j^{\rightarrow}]\right)\right)} \tag{3}$$

In the equation, '· T' represents the transposition operation, and '||' denotes the concatenation operation. After obtaining the normalized attention coefficients, they are utilized to compute a linear combination of the features associated with each coefficient. This linear combination serves as the final output feature for every node. The attention mechanism used in our model is denoted as a (Wh_i, Wh_j), where W*h*i and W*h*$_j$ represent the transformed features of nodes i and j, respectively. This mechanism is parameterized by a weight vector belonging to the real number R and utilizes the LeakyReLU activation function.

In the illustration of multi-head attention, node 1 performs attention on its neighborhood. The attention computations are independent and represented by different arrow styles and colors. The model employs K =3 heads for this multi-head attention. The features obtained from each head are aggregated by concatenating or averaging them to obtain the final output feature h1 for node 1. Applying a nonlinearity, σ) in the following Eq 4:

$$\boldsymbol{h}_i^{\rightarrow\prime} = \sigma\left(\sum_{k\epsilon N_i}\alpha_{ij}W\boldsymbol{h}_j^{\rightarrow}\right) \tag{4}$$

To enhance the stability of the learning process in self-attention, we have discovered that employing multi-head attention is beneficial. This involves executing K-independent attention mechanisms to transform the input features. The transformation is carried out using the following Eq 5:

$$head_k = a\_k(Wh_i, Wh_i) \tag{5}$$

Each attention mechanism, denoted as a_k, operates independently. After the transformation, the features obtained from each head are concatenated together, resulting in the following output feature representation in Eq 6:

$$\boldsymbol{h}_i^{\rightarrow\prime} = ||_{k=1}^{K}\sigma\left(\sum_{k\epsilon N_i}\alpha_{ij}^k W^k\boldsymbol{h}_j^{\rightarrow}\right) \tag{6}$$

In the equation, '||' represents concatenation, αk_{ij} represents the normalized attention coefficients computed by the k_{th} attention mechanism (a^k), and W^k represents the weight matrix of the corresponding input linear transformation. It is important to note that in this setting, the final returned output, h, will consist of KF features (rather than F) for each node. Specifically, if we apply multi-head attention on the final (prediction) layer of the network, concatenation is no longer appropriate. Instead, we employ averaging and postpone the application of the final nonlinearity (such as softmax or logistic sigmoid in classification problems). Eq 7: -

$$\boldsymbol{h}_i^{\rightarrow\prime} = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{k\epsilon N_i}\alpha_{ij}^k W^k\boldsymbol{h}_j^{\rightarrow}\right) \tag{7}$$

E. LAYER-WISE WORKING OF GAN MODEL ARCHITECTURE

Layer-wise demonstration of GAN transformation is outlined through the following steps [33]:

F. DATA ACQUISITION LAYER

The proposed system uses information from two datasets, namely DAPT2020 [39] and Edge IIoT [41], which are accessible to the public, to create the GAN model.

G. LINEAR TRANSFORMATION LAYER

The proposed system applies a linear transformation to the Weighted Matrix W function to obtain feature vectors for the nodes. This function, also known as a projection to the one-hot vectors, encodes the node representations while maintaining the structure of the vector space. Eq 8 computes this linear transformation [42]:

$$Y = WX + b \tag{8}$$

H. COMPUTATION OF ATTENTION COEFFICIENTS

The importance of the features of neighboring nodes is determined through the application of Attention Coefficients, which are computed using Eq 9 [43]:

$$e_{ij} = a\left(W\vec{h_i}, W\vec{h_j}\right) \tag{9}$$

Here a is a function that computes the Attention Coefficients and i and j are neighboring nodes.

I. NORMALIZATION OF ATTENTION COEFFICIENTS

The structures of graphs are diverse, resulting in different numbers of neighbors for each node. To ensure that all neighborhoods have a common scale, attention coefficients are normalized. The normalization function is defined by Eq 10 [44]:

$$\alpha_{ij} = \frac{\exp\left(LeakyReLU\left(e_{ij}\right)\right)}{\sum_{k\epsilon N}\exp\left(LeakyReLU\left(e_{ij}\right)\right)} \tag{10}$$

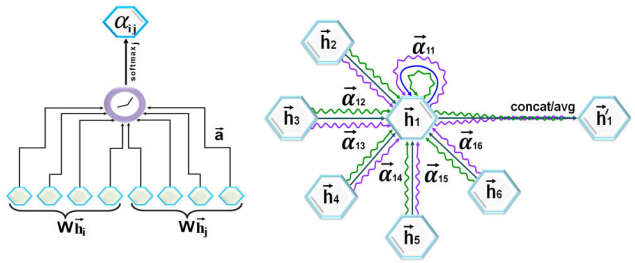

FIGURE 6. Learned features of node.

J. COMPUTATION OF FINAL OUTPUT FEATURES

After normalizing the attention coefficients, the subsequent step is to calculate the group of characteristics connected to the coefficients to achieve the final feature of the network. The method of producing the ultimate output feature is outlined in Eq 11 [45]:-

$$\vec{h_i}' = \sigma\left(\sum_{j\epsilon N}\alpha_{ij}W\vec{h_j}'\right) \tag{11}$$

and its output features are shown in Figure 6.

K. COMPUTATION OF MULTIPLE ATTENTION MECHANISMS

In the last stage, the goal is to enhance the learning process's stability. To accomplish this, Multi-head attention is used to compute several attention maps and obtain a final aggregate of learned representations. This step helps to stabilize the attention process, allowing for multiple independent attention mechanisms to be employed for transforming and concatenating output features. The corresponding equation is Eq 12 [45] defines the Final Learned Feature Computation, whereas its output is shown in Figure 7: -

$$\vec{h_i}' = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j\epsilon N}\alpha_{ij}^{k}W^{k}\vec{h_j}'\right) \tag{12}$$

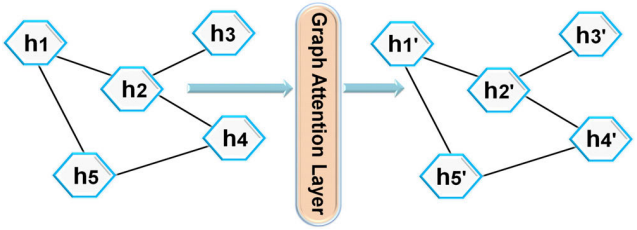

FIGURE 7. Learned features of node.

L. NODE2VEC EMBEDDING LAYER

Node2Vec is a graph embedding technique that converts the nodes in a graph into compact, high-dimensional attribute representations. During the creation of the vector, it takes into account the edges and their weights between the nodes [46]. This technique samples the neighborhood through random walks and trains a hidden layer to predict the probability of one node occurring based on the occurrence of another node, using multiple random neighborhood samples [47]. Node2Vec allows for flexible parameters to explore the graph's neighborhood to obtain rich data representations, ensuring the trade-off between exploration and exploitation, which is essential in graph-optimization problems [47]. Node2Vec is a method used to embed graphs, which converts nodes in a graph into low-dimensional and dense attribute representations. In Node2Vec, each node in the graph acts as an initial point, and a certain number of random walks are generated from these points, forming a structure that serves as input to the Word2Vec model. The objective of training Word2Vec is to exploit the probability of accurately predicting context nodes given the central node. The output of the Word2Vec model is the embedding vectors of predefined sizes that belong to each node in the graph. The embedding

process is illustrated in Figure 8. Five parameters are involved in the Node2Vec embedding process, including the feature embedding size, the number of random walks for each node, the maximum number of nodes visited for each walk, and the p and q parameters, which are used to determine the alpha value [48]. Table 6. provides the hyper-parameters used in the Node2Vec technique. The embedding process is diagrammatically explained in Figure 8.

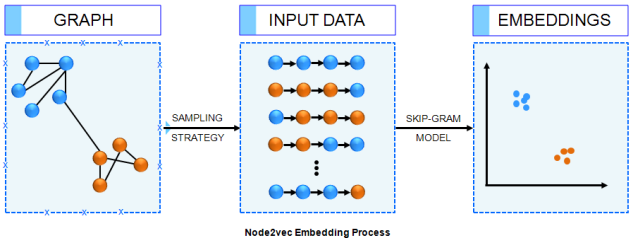

FIGURE 8. NOD2VEC embedding process.

TABLE 6. Hyper-parameters of the Node2Vec model.

Attributes	Numeric
The dimensionality of the feature embedding	50
# Stochastic processes	5
Maximum traversal depth	80
P	0.5
Q	2

M. DETECTION LAYER

The process of detecting APT infiltration in smart I-IoT devices includes constructing a fully connected graph layer through the use of a GAN. This graph layer can identify APT infiltration and categorize data traffic as either APT, Benign, or Non-APT.

N. PERFORMANCE EVALUATION LAYER

The system suggested is assessed by utilizing standard metrics for evaluation, which comprise the detection Accuracy, Precision, Recall, F Score, False Positive Rate, and the duration of Train and Test Execution Time. These metrics are defined as follows: -

TABLE 7. Specifications for hardware/software for proposed framework.

Hardware	Software
GPU 15 GPU	Windows 11 64 Bit and Kali Linux OS
RAM: 16 GB	VMware ver. 16
Graphic Card: 8 GB 1050 Ti	Jupyter Lab and Python 3.9.7 on Anaconda Development framework using the pandas, numpy, Graph Nets, Spectral, TensorFlow, and sklearn modules.

O. MODEL TESTING ACCURACY

Accuracy is a metric used as a reference to evaluate the number of correct predictions made by a classification model, relative to the total number of input samples. The equation provided in Eq 13 best describes how accurately the model is performing [49]:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{13}$$

P. PRECISION, RECALL, F-SCORE, AND FPR

Precision is a metric used to evaluate the accuracy of positive predictions made, specifically for the minority class. It is a benchmark that quantifies the number of correct positive predictions made. Eq 14 and 15 provide the definitions of precision and recall, respectively [50]:-

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

The resulting value of precision is between 0.0, representing no precision, and 1.0, indicating full or ideal precision. The recall is a metric that measures the number of correct positive predictions made out of all possible positive predictions. It indicates the number of missed positive predictions, unlike precision, which identifies the number of accurate positive predictions out of all positive predictions made. Eq 15 can be used to calculate the recall metric.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{15}$$

The result of the recall metric ranges from 0.0, indicating no recall, to 1.0, representing full or ideal recall.

Q. F SCORE

The F-measure or F1 score is a single metric that incorporates both precision and recall, combining the two properties into a single value. It is calculated as the harmonic mean of the two fractions, and Eq 16 can be used to compute the F1 score [51]: -

$$F\ Measure = 2 \times \frac{Precision + Recall}{Precision \times Recall} \tag{16}$$

R. FALSE POSITIVE RATE (FPR)

The accuracy of machine learning can be evaluated using a metric called False Positive Rate (FPR). FPR is calculated as the proportion of negative cases in the dataset that were wrongly classified as positive. Eq 17 provides a description of FPR [52]:-

$$\text{FPR} = \frac{\text{FP}}{(\text{FP} + \text{TN})} \tag{17}$$

S. MODEL TRAIN AND TEST EXECUTION TIME

The measure of a classification system's effectiveness is not solely based on its accuracy in correctly classifying network traffic, but also on its ability to process the data quickly. If processing takes too long, there is a higher chance that APT packets may be missed. In our research, we have evaluated

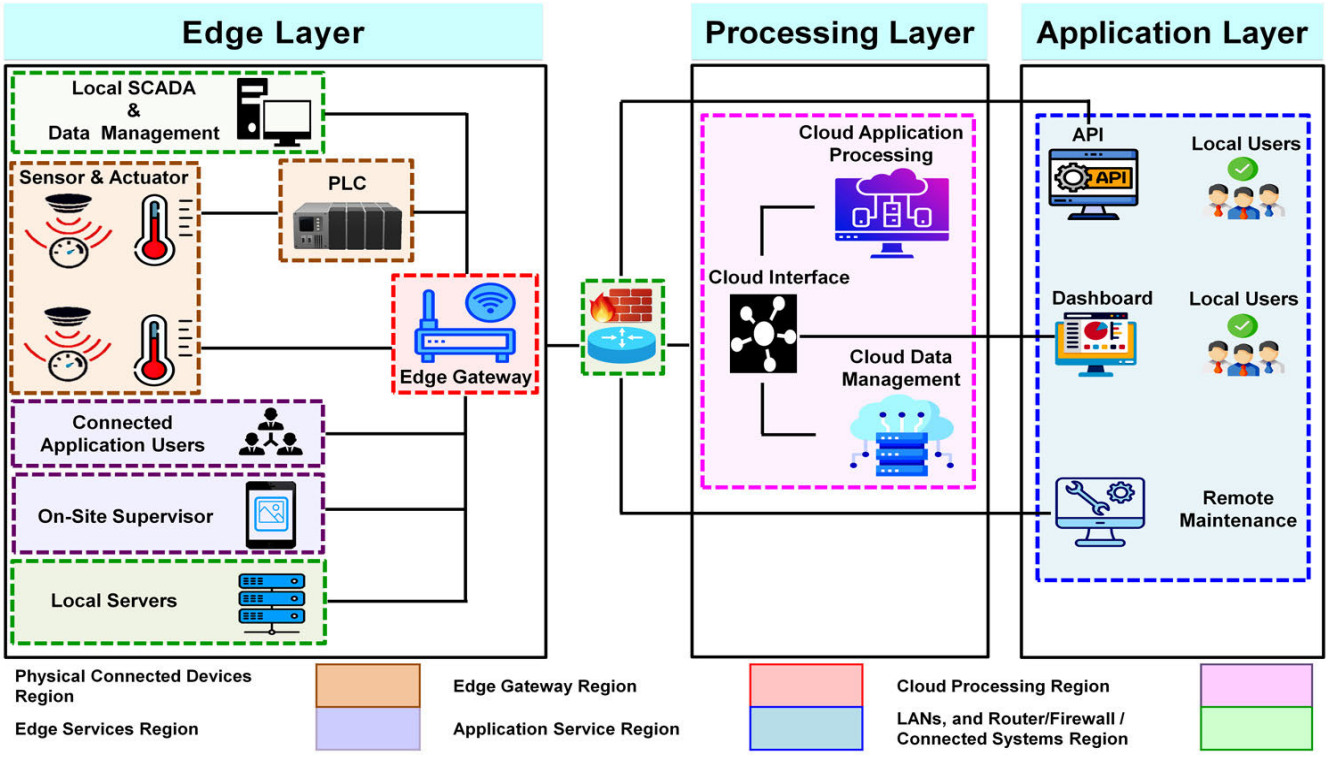

FIGURE 9. Testbed for proposed APT adversarial system.

TABLE 8. Evaluation results: Deep learning algorithms on DAPT2020 malware dataset.

Classification Models	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)	FPR (%)	Train Time (s)	Prediction Time (s)
Proposed	96.97	95.72	96.58	96.58	0.013	10.45	20.56
Signature Based	84.53	82.25	81.35	83.30	20.65	-	-
Layer CNN-1D	90.91	89.54	88.07	90.94	15.789	12.45	11.30
Layers CNN-2D	89.88	85.22	88.5	90.55	14.905	12.78	13.35
Layers CNN-3D	89.24	87.53	90.17	91.59	13.906	13.45	12.33
CNN-LSTM	95.93	94.15	90.15	91.75	2.45	125.39	102.35
Feed Forward Neural Net	89.84	88.34	87.14	87.75	4.67	7.976	6.865
MLP	88.25	85.9	83.20	81.91	4.78	10.78	5.678

the execution times for the training and testing phases of the algorithms that have been applied.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed architecture is evaluated on a system with a standard configuration, and Table 7 shows the hardware and software specifications used for the evaluation of our adversarial framework.

The proposed APT adversarial framework is developed on a windows machines and tested and evaluated on a Linux virtual machine using a generic end-to-end I-IoT security testbed, which simulates real-time I-IoT enabled CPS deployment scenarios [53]. The testbed can be easily modified to support new processes and security scenarios and is a standardized and realistic system for evaluating security solutions in I-IoT networks. It can also analyses I-IoT attack landscapes and provide valuable threat intelligence. Furthermore, the functionality of the proposed testbed is exhibited on various connected devices, communication protocols, and applications. The adversarial APT framework is setup, assessed, and tested on a testbed that simulates I-IoT environment. The testbed layout can be seen in Figure 9.

The wired and wireless devices connected to the testbed are both successfully tested. The performance of the testbed was evaluated by focusing on the edge gateway network and system activities, which are the central points connecting the physical and cyber systems in the I-IoT environment. For the development and evaluation of the proposed system, the

TABLE 9. Evaluation results: DL algorithms on edge-IIoT dataset.

Classification Models	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)	FPR (%)	Train Time (s)	Prediction Time (s)
Proposed	95.25	94.31	99.69	99.43	0.0109	15.01	21.65
Signature Based	89.38	85.88	83.18	81.21	20.96	-	-
Layer CNN-1D	89.03	88.7	87.44	86.93	12.89	11.806	12.122
Layers CNN-2D	90.92	83.79	86.95	89.96	13.805	13.19	12.991
Layers CNN-3D	91.07	88.42	88.76	91.97	12.99	12.94	11.64
CNN-LSTM	95.10	89.25	87.15	88.13	2.987	12.392	6.335
Feed Forward Neural Net	88.26	86.20	85.53	86.12	11.455	8.976	7.865
MLP	85.53	1.09	1.23	1.001	3.67	9.786	4.678

TABLE 10. Input parameters for random decision forest.

Hyper-Parameters	Value
Max_depth	5,10,100,500,1000
Min_samples_split	5,10,100,500
N_estimators	5,10,50,100,500
Criterion	Gini
Class_weight	Balanced
Cross_Validation (CV)	3
N_jobs	-1

TABLE 11. Input parameters for decision tree.

Hyper-Parameters	Value
Max_depth	5,10,20,50,100,500
Min_samples_split	5,10,100,500
N_estimators	5,10,50,100,500
Criterion	Gini
Splitter	Best
Class_weight	Balanced
Cross_Validation (CV)	3
N_jobs	1

TABLE 12. Input parameters for SVM.

Hyper-Parameters	Value
Alpha	$10^{-8} - 10^{3}$
Penalty	L1, L2
Loss	Hinge
Cross-validation (CV)	5

TABLE 13. Input parameters for logistic regression.

Hyper-Parameters	Value
Alpha	$10^{-8} - 10^{3}$
Penalty	L1, L2
Loss	Hinge
Class_weight	Balanced
Cross-validation (CV)	5
N_jobs	-1

TABLE 14. Input parameters for GNB.

Hyper-Parameters	Value
Var_smoothing	$10^{-9} - 10^{3}$
Cross-validation (CV)	5
N_jobs	-1

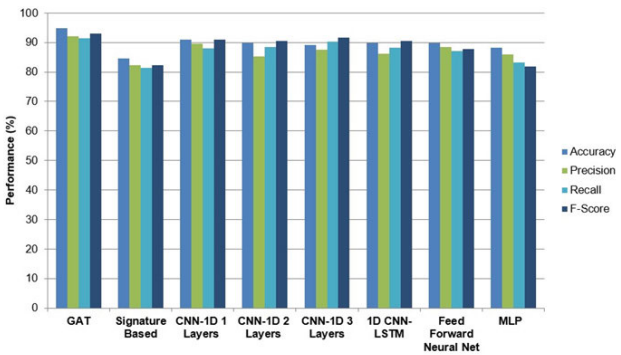

FIGURE 10. Comparison results of GAN with DNN algorithms - DAPT2020 malware dataset.

Pytorch Geometric module of the Pytorch framework is used to build the GAN [54]. The Pytorch framework offers temporal geometric DL for researchers and ML practitioners in a user-friendly and integrated environment, making it a better option compared to other DL libraries. During the training stage, different experimental configurations were applied to the GAN models of the proposed system by varying the number of hidden units in the convolutional layers. Specifically,

TABLE 15. Evaluation results: Conventional ML algorithms on DAPT2020 dataset.

Classification Models	Accuracy	Precision	Recall	F-Score	FPR	Train Time (s)	Prediction Time (s)
Random Decision Forest	87.39	88.67	87.76	87.67	15.43	207.56	156.67
Support Vectors Machine	62.32	66.45	67.45	65.75	20.44	66.42	110
Decision Tree	84.54	83.54	82.36	87.99	14.67	155	145
Naives Bayes	75.73	74.38	73.45	76.58	13.78	140	135
Logistic Regression	78.45	77.89	76.78	75.45	19.67	255	245

TABLE 16. Evaluation results: Conventional ML algorithms on edge IIoT dataset.

Classification Models	Accuracy	Precision	Recall	F-Score	FPR	Train Time (s)	Prediction Time (s)
Random Decision Forest	80.96	78.09	77.56	76.29	14.54	91	76
Support Vectors Machine	63.4	62.25	61.31	63.001	18.44	96	85
Decision Tree	72.47	71.63	72.32	73.45	15.65	153	146
Naives Bayes	61.67	64.99	63.32	63.45	17.21	191	184
Logistic Regression	76.35	75.48	72.36	73.31	18.64	289	286

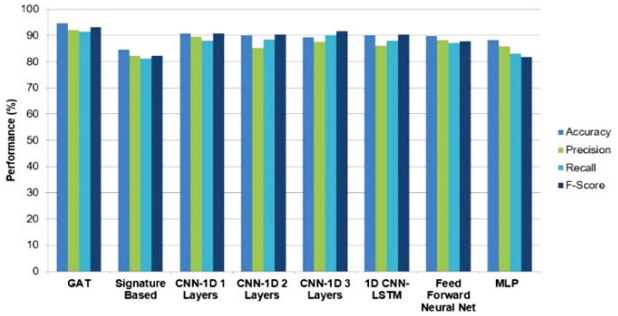

FIGURE 11. Comparison results of GAN with DNN algorithms – Edge-IIoT dataset.

the experimental setup included hidden units of 16, 32, 54, and 128. To assess the statistical significance of these experimental setups, the Wilcoxon Signed Rank Test [55] is used with a significance level of 0.05. The results obtained from applying various hidden units on GAN with Node2Vec embedding are shown in Tables 8-9 on the DAPT2020 Malware and Edge-IIoT datasets respectively. A graphical comparison of the results is shown in Figures 10-11.

We evaluated our models on the DAPT2020 and edge IIoT datasets, and our adversarial models perform exceptionally well in terms of accuracy, recall, and precision with very low FPR (0.013%). The results demonstrate that the GAN algorithm stands out with the highest accuracy rate of 96.97% when using 128 hidden units in its convolutional layers on the DAPT2020 malware dataset. Moreover, it takes about 10.45s to compile the model and 20.56s to predict the results. The same model also achieves a prediction accuracy of 95.97% with FPR of 0.0109% and train and prediction time of 15.01s and 21.65s to detect APT assaults respectively on the Edge IIoT dataset when using the same hidden parameters. Although the CNN-LSTM model also obtains excellent prediction accuracy of 95.93% (Accuracy, Precision, Recall, F1-score) on the DAPT2020 dataset, however, its FPR stands high at 2.45% with the model requiring 125s during the training phase and 102s during the prediction phase on DAPT2020 datasets. Likewise, the outcomes of the model on the Edge-IIoT dataset are almost indistinguishable.

A. EXPERIMENTATION WITH CONVENTIONAL ML ALGORITHMS

Conventional ML algorithms for classification problems have been developed and compared to the proposed DL approach. A similar structure of DL was used for conventional classification to make an objective evaluation and compare its results with other classifier algorithms. Standard data transformation techniques, application of the classification algorithm, and the evaluation of results were carried out, including nominalization and elimination of unnecessary features. The algorithms tested were LinearSVC with 500 iterations, RandomForestClassifier, DecisionTreeClassifier from sklearn, GaussianNB from NumPy, and LogisticRegression from the Sklearn ML libraries. Standard evaluation metrics such as Accuracy, Precision, Recall, and F-score are used to evaluate the traditional ML algorithms. Input parameters of selected algorithms are provided in the following Table 10-14.

The experimental results showed the highest accuracy of 87.39%, a precision of 88.67%, recall of 87.76%, f1-score, and 87.32% respectively, with the RDF algorithm outperforming other conventional algorithms in the DAPT2020 malware and Edge-IIoT datasets. the results of the traditional MLalgorithms using the same datasets are shown in Tables 15-16.

In summary, based on the evaluation results, it can be concluded that GAN performs better than other ML algorithms in accurately classifying complex malware signatures, including

TABLE 17. Comparison of proposed methodology with existing state-of-the-art.

Papers	Techniques	Datasets	Accuracy
Kim et al. [13]	DL-based malware detection system using CNN	Malimg dataset	The approach achieves a very high accuracy of 98.93% which needs to be tested on a real-time APT dataset.
Huang et al. [14]	Game-theoretic framework to design proactive and cross-layer defenses for cyber-physical systems in a holistic manner against APT in autonomous system	-	The approach achieves a payoff of 56% higher than a primitive defense mechanism.
Ghafir et al. [12]	Autonomous APT Detection Approach in generalized domain	Customized ML Dataset	Achieves an accuracy of 84.8%
Wang et al. [30]	Autonomous APT Detection system for detection of Command and Control (C & C) communication	Los Almos Laboratory APT Dataset	Achieves 83.3% TPR.
Yu et al.[23]	Bidirectional encoder representations from transformers (BERT) scheme that detects APT attack sequences.	Uses five simulated attack categories: "NORMAL", "PROBE", "DOS", "U2R" and "R2L".	The scheme provides an accuracy of 99%.
Siniosoglou et al.[24]	MENSA (Anomaly Detection and classification) adopts a novel Auto-Encoder-Generative Adversarial Network (GAN) architecture	Modbus/TCP network flows, DNP3 network flows, and operational time-series electricity measurements data	TRP and FPR of 0.947%, 0.812 and 0.036 respectively
Kumar et al. [25]	The connected phases of the attack are employed to create a concise and comprehensive APT Campaign Graph, which can be used to monitor the advancement of the APT campaign and apply suitable measures to counter it.	The approach uses a synthetic APT campaign injected into the CSE-IDC2018 intrusion detection dataset	Precision Accuracy of 0.996%
Du et al. [56]	DL-based NIDS using the CNN-LSTM model is presented for the wireless sensing scenario of the I-IoT to effectively distinguish and identify network traffic data.	KDDCUP99, NSL_KDD and UNSW_NB15	The approach achieves very high detection accuracy (0.944-0.999) on all 3 datasets, however, the approach needs to be validated on pure APT datasets.
Hnamte et al.[57]	Two-stage DL technique hybridizing LSTM and Auto-Encoders (AE) technique has been applied for detecting attacks.	CICIDS2017 and CSE-CICDIS2018	The approach achieves an accuracy of 99.99% and 99.10% on subject datasets. However, similar results on the APT dataset need to be ascertained in the I-IoT domain.
Halbouni et al.[58]	Dynamic line graph neural network (DLGNN)-based IDS method with semisupervised learning has been applied. The model converts network traffic into a series of spatiotemporal graphs	CIC-IDS 2017, UNSW-NB15, and WSN-DS	The approach achieves an accuracy of more than 98% on all datasets. The results need to be validated on pure APT datasets on sensor networks.
Duan et al. [59]	Dynamic line graph neural network (DLGNN)-based IDS method with semisupervised learning has been applied. The model converts network traffic into a series of spatiotemporal graphs	NIDS	The approach achieves 98.15–99.8% accuracy in abnormality detection with fewer labeled samples.
Hnamte et al[32]	An intelligent and efficient DCNNBiLSTM system based on combination of CNN and LSTM model is proposed, for the detection of attacks. The model is trained on real-time traffic datasets and achieves high-end accuracy on network traffic.	CICIDS2018 and Edge_IIoT datasets.	Achieves very high accuracy of 100% and 99.64% on relevant datasets, however, its efficacy in terms of performance it needs to be validated for the I-IoT domain.
Proposed APT Adversarial System	Applies Graph Attention Network (GAN)	DAPT2020 and Edge IIoT datasets	96.97% accuracy with a 20.56s prediction time on DAPT2020 and 95.97% with a 21.65s prediction time on Edge I-IoT datasets, respectively.

APT data signatures. A comparative analysis with other techniques for APT detection and classification is presented in Table 17, which demonstrates the superior performance of the proposed GAN approach in terms of accuracy for detection and classification with minimal processing time, well suited for I-IoT enabled CPS domain. Various researchers in the

field have proposed ML and independent system approaches that exhibit exceptional detection rates. For example, Kumar and Thing [25] and Hnamte et al. [57] presented an independent and DL based methods that achieved high end accuracy of 0.99% and 100% prediction rate, respectively. However, their scope is restricted to a small set of APT attack phases, and their effectiveness is assessed using a simulated dataset that may not provide a precise representation of APT characteristics in IoT-enabled CPS systems. On the other hand, the adversarial method outlined in the aforementioned tables requires minimal computation time for processing APT data packets, leading to better overall performance by reducing the duration of training and prediction and minimizing computation power consumption during all phases of APT attacks. Furthermore, scalability can be further enhanced and it can be quickly deployed. Additionally, the system is built in a modular and portable way increasing its flexibility and adaptability.

VI. CONCLUSION AND FUTURE WORK

I-IoT-enabled CPSs control the infrastructure in our society. In recent years, there have been instances of vulnerabilities being exploited in these mission-critical systems, which have a larger attack surface compared to traditional IT systems. The poor security measures of different I-IoT devices and prevalent conditions make them vulnerable to APT attacks.

The current security systems in I-IoT-enabled CPSs are keeping pace with technological advancements, but they are not effective and far from adequate in protecting sensitive systems. This study presents a DL-based APT campaign detection system for I-IoT-enabled CPSs that uses GAN algorithms for APT attack detection and compensation. The experiment results show that the GAN approach is most suitable for a mission-critical robust system of I-IoT-enabled CPSs and effectively detects complex APT attacks with a prediction accuracy and time of 96.97% (20.56s) and 95.25% (21.65s) on the DAPT2020 malware and Edge I-IoT datasets respectively. The comparison of the proposed approach with the state of the art suggests that the DL approach is superior in detecting complex APT malware in the hazardous domain of I-IoT-enabled CPSs. Overall, the proposed DL-based approach is superior in detecting complex APT malware in the hazardous domain of I-IoT-enabled CPSs. Future and further aspects that could be explored include the viability of other NN variants, such as RNN, MLP, GRU, etc. in the subject domain with higher computational resources for better performance in detection. Secondly, leveraging the attention mechanism to conduct a comprehensive analysis of model interpretability. Finally, incorporating edge features into the model would enable us to address a wide range of problems by capturing relationships among nodes in the subject domain can also be explored.

REFERENCES

[1] D. Huang, C. Lin, C. Chen, and J. Sze, "The internet technology for defect detection system with deep learning method in smart factory," in *Proc. 4th Int. Conf. Inf. Manag. (ICIM)*, May 2018, pp. 98–102.

[2] E. Inshakova, A. Inshakova, and A. Goncharov, "Engineered nanomaterials for energy sector: Market trends, modern applications and future prospects," in *Proc. IOP Conf. Ser., Mater. Sci. Eng.*, 2020, Art. no. 032031.

[3] S. Al-Sarawi, M. Anbar, R. Abdullah, and A. B. Al Hawari, "Internet of Things market analysis forecasts, 2020–2030," in *Proc. 4th World Conf. Smart Trends Syst., Secur. Sustainability (WorldS)*, Jul. 2020, pp. 449–453.

[4] S. Bhattacharjee, *Practical Industrial Internet of Things Security: A Practitioner's Guide to Securing Connected Industries*. Birmingham, U.K.: Packt, 2018.

[5] M. B. Ahmad, A. Akram, M. Asif, and S. Ur-Rehman, "Using genetic algorithm to minimize false alarms in insider threats detection of information misuse in windows environment," *Math. Problems Eng.*, vol. 2014, pp. 1–12, Jan. 2014.

[6] A. Salifu, "The impact of internet crime on development," *J. Financial Crime*, vol. 15, no. 4, pp. 432–443, Oct. 2008.

[7] S. A. Amro, D. A. Elizondo, A. Solanas, and A. Martinez-Balleste, "Evolutionary computation in computer security and forensics: An overview," in *Computational Intelligence for Privacy and Security*. Springer, 2022, pp. 25–34.

[8] M. Asif, Z. Aziz, M. B. Ahmad, A. Khalid, H. A. Waris, and A. Gilani, "Blockchain-based authentication and trust management mechanism for smart cities," *Sensors*, vol. 22, no. 7, p. 2604, Mar. 2022.

[9] M. M. Salim, S. Rathore, and J. H. Park, "Distributed denial of service attacks and its defenses in IoT: A survey," *J. Supercomput.*, vol. 76, no. 7, pp. 5320–5363, Jul. 2020.

[10] S. H. Javed, M. B. Ahmad, M. Asif, S. H. Almotiri, K. Masood, and M. A. A. Ghamdi, "An intelligent system to detect advanced persistent threats in industrial Internet of Things (I-IoT)," *Electronics*, vol. 11, no. 5, p. 742, Feb. 2022.

[11] L. Ballerini, O. Cordon, S. Damas, J. Santamaría, I. Aleman, and M. Botella, "Craniofacial superimposition in forensic identification using genetic algorithms," in *Proc. 3rd Int. Symp. Inf. Assurance Secur.*, Aug. 2007, pp. 429–434.

[12] I. Ghafir, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, and F. J. Aparicio-Navarro, "Detection of advanced persistent threat using machine-learning correlation analysis," *Future Gener. Comput. Syst.*, vol. 89, pp. 349–359, Dec. 2018.

[13] M. Azizjon, A. Jumabek, and W. Kim, "1D CNN based network intrusion detection with normalization on imbalanced data," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Feb. 2020, pp. 218–224.

[14] L. Huang and Q. Zhu, "A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems," *Comput. Secur.*, vol. 89, Feb. 2020, Art. no. 101660.

[15] S. Tamy, H. Belhadaoui, M. A. Rabbah, N. Rabbah, and M. Rifi, "An evaluation of machine learning algorithms to detect attacks in SCADA network," in *Proc. 7th Medit. Congr. Telecommun. (CMT)*, Oct. 2019, pp. 1–5.

[16] J. Qian, X. Du, B. Chen, B. Qu, K. Zeng, and J. Liu, "Cyber-physical integrated intrusion detection scheme in SCADA system of process manufacturing industry," *IEEE Access*, vol. 8, pp. 147471–147481, 2020.

[17] J. Gao, L. Gan, F. Buschendorf, L. Zhang, H. Liu, P. Li, X. Dong, and T. Lu, "LSTM for SCADA intrusion detection," in *Proc. IEEE Pacific Rim Conf. Commun., Comput. Signal Process. (PACRIM)*, Aug. 2019, pp. 1–5.

[18] B. Stewart, L. Rosa, L. A. Maglaras, T. J. Cruz, M. A. Ferrag, P. Simoes, and H. Janicke, "A novel intrusion detection mechanism for SCADA systems which automatically adapts to network topology changes," *EAI Endorsed Trans. Ind. Netw. Intell. Syst.*, vol. 4, no. 10, Feb. 2017, Art. no. 152155.

[19] G. George and S. M. Thampi, "A graph-based security framework for securing industrial IoT networks from vulnerability exploitations," *IEEE Access*, vol. 6, pp. 43586–43601, 2018.

[20] M. Teixeira, T. Salman, M. Zolanvari, R. Jain, N. Meskin, and M. Samaka, "SCADA system testbed for cybersecurity research using machine learning approach," *Future Internet*, vol. 10, no. 8, p. 76, Aug. 2018.

[21] C.-Y. Lin and S. Nadjm-Tehrani, "Timing patterns and correlations in spontaneous SCADA traffic for anomaly detection," in *Proc. RAID*, 2019, pp. 73–88.

[22] X. Zhou, Z. Xu, L. Wang, K. Chen, C. Chen, and W. Zhang, "Behavior based anomaly detection model in SCADA system," in *Proc. MATEC Web Conf.*, 2018, p. 01011.

[23] K. Yu, L. Tan, S. Mumtaz, S. Al-Rubaye, A. Al-Dulaimi, A. K. Bashir, and F. A. Khan, "Securing critical infrastructures: Deep-learning-based threat detection in IIoT," *IEEE Commun. Mag.*, vol. 59, no. 10, pp. 76–82, Oct. 2021.

[24] I. Siniosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Fouliras, and P. Sarigiannidis, "A unified deep learning anomaly detection and classification approach for smart grid environments," *IEEE Trans. Netw. Service Manag.*, vol. 18, no. 2, pp. 1137–1151, Jun. 2021.

[25] A. Kumar and V. L. L. Thing, "RAPTOR: Advanced persistent threat detection in industrial IoT via attack stage correlation," 2023, *arXiv:2301.11524*.

[26] H. Yao, P. Gao, P. Zhang, J. Wang, C. Jiang, and L. Lu, "Hybrid intrusion detection system for edge-based IIoT relying on machine-learning-aided detection," *IEEE Netw.*, vol. 33, no. 5, pp. 75–81, Sep. 2019.

[27] X. Xiao, Z. Tang, C. Li, B. Xiao, and K. Li, "SCA: Sybil-based collusion attacks of IIoT data poisoning in federated learning," *IEEE Trans. Ind. Informat.*, vol. 19, no. 3, pp. 2608–2618, Mar. 2023.

[28] S. Dhelim, N. Aung, M. T. Kechadi, H. Ning, L. Chen, and A. Lakas, "Trust2Vec: Large-scale IoT trust management system based on signed network embeddings," *IEEE Internet Things J.*, vol. 10, no. 1, pp. 553–562, Jan. 2023.

[29] I. Ghafir, K. G. Kyriakopoulos, S. Lambotharan, F. J. Aparicio-Navarro, B. Assadhan, H. Binsalleeh, and D. M. Diab, "Hidden Markov models and alert correlations for the prediction of advanced persistent threats," *IEEE Access*, vol. 7, pp. 99508–99520, 2019.

[30] X. Wang, K. Zheng, X. Niu, B. Wu, and C. Wu, "Detection of command and control in advanced persistent threat based on independent access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[31] J. V. Chandra, N. Challa, and S. K. Pasupuleti, "A practical approach to e-mail spam filters to protect data from advanced persistent threat," in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Mar. 2016, pp. 1–5.

[32] V. Hnamte and J. Hussain, "DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system," *Telematics Informat. Rep.*, vol. 10, Jun. 2023, Art. no. 100053.

[33] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," 2017, *arXiv:1710.10903*.

[34] J. You, J. Leskovec, K. He, and S. Xie, "Graph structure of neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 10881–10891.

[35] A. Tato and R. Nkambou, "Improving Adam optimizer," in *Proc. ICLR Workshop Submissions*, May 2023.

[36] S. Barnum, "Standardizing cyber threat intelligence information with the structured threat information expression (STIX)," *Mitre Corp.*, vol. 11, pp. 1–22, Jan. 2012.

[37] H. Haddadpajouh, A. Azmoodeh, A. Dehghantanha, and R. M. Parizi, "MVFCC: A multi-view fuzzy consensus clustering model for malware threat attribution," *IEEE Access*, vol. 8, pp. 139188–139198, 2020.

[38] H. HaddadPajouh, A. Dehghantanha, R. Khayami, and K.-K.-R. Choo, "A deep recurrent neural network based approach for Internet of Things malware threat hunting," *Future Gener. Comput. Syst.*, vol. 85, pp. 88–96, Aug. 2018.

[39] S. Myneni, A. Chowdhary, A. Sabur, S. Sengupta, G. Agrawal, D. Huang, and M. Kang, "DAPT 2020-constructing a benchmark dataset for advanced persistent threats," in *Proc. Int. Workshop Deployable Mach. Learn. Secur. Defense*, San Diego, CA, USA, Aug. 2020, pp. 138–163.

[40] B. Chen, J. Wan, Y. Lan, M. Imran, D. Li, and N. Guizani, "Improving cognitive ability of edge intelligent IIoT through machine learning," *IEEE Netw.*, vol. 33, no. 5, pp. 61–67, Sep. 2019.

[41] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning," *IEEE Access*, vol. 10, pp. 40281–40306, 2022.

[42] S. K. Maurya, X. Liu, and T. Murata, "Improving graph neural networks with simple architecture design," 2021, *arXiv:2105.07634*.

[43] D. Buterez, I. Bica, I. Tariq, H. Andrés-Terré, and P. Lio, "CellVGAE: An unsupervised scRNA-seq analysis workflow with graph attention networks," *Bioinformatics*, vol. 38, no. 5, pp. 1277–1286, Feb. 2022.

[44] L. Gong and Q. Cheng, "Exploiting edge features for graph neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 9203–9211.

[45] K. Wang, W. Shen, Y. Yang, X. Quan, and R. Wang, "Relational graph attention network for aspect-based sentiment analysis," 2020, *arXiv:2004.12362*.

[46] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, "Heterogeneous graph neural networks for malicious account detection," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manag.*, Oct. 2018, pp. 2077–2085.

[47] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 855–864.

[48] S. De Winter, T. Decuypere, S. Mitrovic, B. Baesens, and J. De Weerdt, "Combining temporal aspects of dynamic networks with Node2Vec for a more efficient dynamic link prediction," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2018, pp. 1234–1241.

[49] B. Juba and H. S. Le, "Precision-recall versus accuracy and the role of large data sets," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 4039–4048.

[50] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, *arXiv:2010.16061*.

[51] E. G. Dada, J. S. Bassi, Y. J. Hurcha, and A. H. Alkali, "Performance evaluation of machine learning algorithms for detection and prevention of malware attacks," *IOSR J. Comput. Eng.*, vol. 21, pp. 18–27, May 2019.

[52] P. Wang, Q. Wu, J. Cao, C. Shen, L. Gao, and A. V. D. Hengel, "Neighbourhood watch: Referring expression comprehension via language-guided graph attention networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 1960–1968.

[53] M. Al-Hawawreh and E. Sitnikova, "Developing a security testbed for industrial Internet of Things," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5558–5573, Apr. 2021.

[54] P. Mishra, "Introduction to neural networks using PyTorch," in *PyTorch Recipes*. Berlin, Germany: Springer, 2022, pp. 117–133.

[55] R. F. Woolson, "Wilcoxon signed-rank test," in *Wiley Encyclopedia of Clinical Trials*. Wiley, 2007, pp. 1–3.

[56] J. Du, K. Yang, Y. Hu, and L. Jiang, "NIDS-CNNLSTM: Network intrusion detection classification model based on deep learning," *IEEE Access*, vol. 11, pp. 24808–24821, 2023.

[57] V. Hnamte, H. Nhung-Nguyen, J. Hussain, and Y. Hwa-Kim, "A novel two-stage deep learning model for network intrusion detection: LSTM-AE," *IEEE Access*, vol. 11, pp. 37131–37148, 2023.

[58] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, "CNN-LSTM: Hybrid deep neural network for network intrusion detection system," *IEEE Access*, vol. 10, pp. 99837–99849, 2022.

[59] G. Duan, H. Lv, H. Wang, and G. Feng, "Application of a dynamic line graph neural network for intrusion detection with semisupervised learning," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 699–714, 2023.

SAFDAR HUSSAIN JAVED is currently pursuing the Ph.D. degree with the College of Computing and Information Sciences, Karachi Institute of Economics and Technology (KIET), Karachi, Pakistan. His current research interest includes cyber security in combination with machine and deep learning systems.

MAAZ BIN AHMAD received the Ph.D. degree in computer engineering from the Centre for Advanced Studies in Engineering (CASE), Islamabad, Pakistan, in 2014. Currently, he is an Associate Professor with the College of Computing and Information Sciences, Karachi Institute of Economics and Technology, Karachi, Pakistan. He has published more than 35 research articles. His current research interests include network security, image and video processing, machine learning, and multimedia.

MUHAMMAD ASIF received the Ph.D. degree in digital image processing from the Capital University of Science and Technology (CUST), Islamabad, Pakistan, in 2016. Currently, he is an Associate Professor with the Department of Computer Science, Lahore Garrison University (LGU), Lahore, Pakistan. He has contributed more than 45 research articles. His current research interests include image and video processing, computer vision, machine learning, parallel processing, embedded system optimization, and network security.

WASEEM AKRAM received the M.S. degree in computer science from The Islamia University of Bahawalpur, Pakistan, in 2020. He is currently with Lahore Garrison University, Lahore. His current research interests include network security, healthcare authentication, and authenticated key agreement protocols using lightweight cryptography.

KHALID MAHMOOD (Senior Member, IEEE) received the Ph.D. degree in computer science from International Islamic University, Islamabad, Pakistan, in 2018. He is currently with the National Yunlin University of Science and Technology, Yunlin, Taiwan. His current research interests include lightweight cryptography, design, and development of authentication protocols using lightweight cryptographic solutions for diverse infrastructures.

ASHOK KUMAR DAS (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering, the M.Tech. degree in computer science and data processing, and the M.Sc. degree in mathematics from IIT Kharagpur, India. He is currently an Associate Professor with the Center for Security, Theory and Algorithmic Research, IIIT Hyderabad, India, and a Visiting Faculty with the Virginia Modeling, Analysis and Simulation Center, Old Dominion University, Suffolk, VA, USA. His Google Scholar H-index is 77 and his I10-index is 220 with more than 16,960 citations. His current research interests include cryptography, systems, and network security, including security in smart grids, the Internet of Things (IoT), the Internet of Drones (IoD), the Internet of Vehicles (IoV), cyber-physical systems (CPS) and cloud computing, intrusion detection, blockchain, AI/ML security, and post-quantum cryptography. He has authored more than 350 papers in international journals and conferences in the above areas, including more than 300 reputed journal articles. He was a recipient of the Institute Silver Medal from IIT Kharagpur. He has been listed in the Web of Science (Clarivate^{TM}) Highly Cited Researcher in recognition of his exceptional research performance, in 2022. He served as the Technical Program Committee Chair for the First International Congress on Blockchain and Applications (BLOCKCHAIN'19), Avila, Spain, in June 2019; the International Conference on Applied Soft Computing and Communication Networks (ACN'20), Chennai, India, in October 2020; and the Second International Congress on Blockchain and Applications (BLOCKCHAIN'20), L'Aquila, Italy, in October 2020. He was/is on the editorial board of IEEE Systems Journal, *Journal of Network and Computer Applications* (Elsevier), *Computer Communications* (Elsevier), *Journal of Cloud Computing* (Springer), *Cyber Security and Applications* (Elsevier), *IET Communications*, *KSII Transactions on Internet and Information Systems*, and *International Journal of Internet Technology and Secured Transactions* (Inderscience).

SACHIN SHETTY (Senior Member, IEEE) received the Ph.D. degree in modeling and simulation from Old Dominion University, in 2007. He was an Associate Professor with the Electrical and Computer Engineering Department, Tennessee State University, USA. He is currently a Professor with the Virginia Modeling, Analysis and Simulation Center, Old Dominion University. He holds a joint appointment with the Center for Cybersecurity Education and Research and the Department of Modeling, Simulation and Visualization Engineering. He has authored or coauthored more than 200 research papers in journals and conference proceedings and two books. His current research interests include the intersection of computer networking, network security, and machine learning.