A Review on Emerging Machine Learning Models for IDS in IoT

MD SHAHEDUL HAQUE[*] and AMIRREZA GHAFOORI[*], Departments of CS & ECE, Virginia Polytechnic Institute and State University, USA

Abstract: The Internet of Things (IoT) is a system for connecting objects from wide ranges such as common, day to day use physical devices to sophisticated industrial tools. With the proliferation of IoT devices, it has drawn researchers' attention toward its security since they can be easily comprised. The system of Internet of Things (IoT) is vulnerable to various cyber attacks than desktop computers which hampers critical and real-time applications. To secure any IoT system, it is needed that an anomaly is detected as soon as possible so that the system does not halt due to the attack. Another aspect of intrusion detection in IoT system is handing the compromised component to mitigate the impact of the attack. Several machine learning models with various capability has been deployed to address these issues in IoT security by building intrusion detection systems (IDS). This paper provides a detailed discussion on attacks against IoT networks and emerging machine learning models applied in this scenario.

CCS Concepts: • **Security and privacy** → **Intrusion/anomaly detection and malware mitigation**; **Intrusion detection systems**; **Artificial immune systems**.

Additional Key Words and Phrases: IoT, Industrial IoT, Cyber-Physical Systems (CPS), machine learning, active attack, passive attack, prevention, intrusion detection system, anomaly detection, mitigation.

ACM Reference Format:
Md Shahedul Haque and Amirreza Ghafoori. . A Review on Emerging Machine Learning Models for IDS in IoT. In *FIS final project report, May, 2023, Arlington, VA.* ACM, New York, NY, USA, 15 pages.

1 INTRODUCTION

Security of IoT systems has been a hot research topic due to the increasing use of IoT devices in every sphere of our life. All IoT devices can be categorized into two broad categories: Switch (send commands) and Sensor (Collects data and sends it across to some other device). IoT gateways or edge devices pave the way for the connection for sensors. This distributed system is prone to several active and passive cyber attacks such as eavesdropping, traffic analysis, jamming, flooding, DDoS and so on. To detect such attack, two approaches have been adopted - signature based intrusion detection and anomaly based intrusion detection. In this work, we will focus on the contribution of recent advancement in machine learning models for the intrusion detection of IoT systems. In their survey [2] on security of the Internet of Things, they discuss the vulnerabilities, attack and their countermeasures as defense strategy against those attacks. According to their findings, most of the proposed detection or prevention and mitigation methods are on the network layer. Similar observation can be found in [14]. For defending IoT systems, [15] integrates Complex Event Processing (CEP) technology and the Machine Learning (ML) paradigm. Other approaches have utilized Deep Recurrent Neural Networks based detection [16]. For deploying distributed detection system, Federated Learning based anomaly detection [11] is quite effective. [6] provides an alternative solution using generative adversarial network. [3] provides a study about intrusion detection systems methods for IoT based on machine learning. Also, [17] provides a detailed review on IDS for IoT devices while [9] provides a survey on deep learning based IDS for automotive applications. On the other hand, [5] utilizes GAN to build an IDS for 5G Enabled Future Metaverse. Also there has been approached

[*] Both authors contributed equally to this research.

towards combining more than one machine learning models to build the IDS such as [19], [20], [21], [13] etc.

This survey aims to provide a comprehensive overview of the applications of advanced machine learning models for the security of IoT devices namely the intrusion detection systems. We address the following research questions (RQs):

- **RQ1: What are the common attacks in IoT systems?**
  The attacks can be divided into two parts - passive attack and active attack. Most common passive attacks are eavesdropping, traffic analysis, etc while the most common active attacks are jamming, flooding, DoS, DDOS, replay attack, hole attack, Sybil types, etc.
- **RQ2: What are the application domains of IDS in IoT system with emerging machine learning models?**
  Several noteworthy application domains of IDS in IoT system with machine learning are wireless sensor networks, cloud environment, enterprise network, smart grid technologies, transportation, smart home, smart cities, healthcare applications, etc.
- **RQ3: Which ML based architecture is used most often in IDS for IoT systems?**
  Deep learning architectures, decision trees, random forests, support vector machines, and Naive Bayes are commonly used machine learning-based architectures for intrusion detection systems in IoT systems. The specific architecture used depends on the system's requirements and constraints, such as available computational resources, data characteristics, and desired performance metrics.
- **RQ4: Which type of data instance and datasets are most commonly used for ML based IDS in IoT systems?**
  ML-based IDS in IoT systems commonly use network traffic data, system logs, botnet and malware data, and IoT-specific datasets for training and testing ML models. The choice of data depends on the specific characteristics of the system and the targeted attacks.
- **RQ5: Which metrics are used to evaluate the performance of ML models in generating data and preserving the safety of IoT systems?**
  Metrics commonly used to evaluate ML models for generating data and preserving the safety of IoT systems include accuracy, precision and recall, F1-score, false positive rate and false negative rate, AUC, mean squared error, and computational time. The choice of metrics depends on the specific requirements of the system and the type of data being analyzed.
- **RQ6: Which attack prevention and/or detection techniques are used along with ML models?**
  Strong password, antivirus, firewalls, regular update, monitoring for anomaly detection, etc.

The difference between this study and previous works is summarized below:

(1) Emerging Machine learning based models for IoT intrusion detection
(2) Discussion on recent changes and challenges in IoT system's intrusion detection

The remainder of this paper is organized as follows. The related works are discussed in Section 2. Sections 3 discloses the existing attacks in IoT systems. Sections 4, 5, and 6 introduce various machine learning models from the algorithm, methodology, and application perspectives. Defense techniques are discussed in section 7. Tables 1 and 2 list the interesting machine learning algorithms and application fields, which are discussed in Sections 4 and 6, respectively. Finally, Section 8 concludes the survey.

2 RELATED WORKS

IoT systems is vulnerable to numerous attacks in the IoT network due to their resource constraints and variation of device types. Intrusion detection system is a fundamental component in cyber-physical systems to prevent cyber-attacks of various kinds. Machine learning models have gained popularity in intrusion detection system due to their significant ability in handling large amount of data and explicitly or implicitly learning from the data. Main security issues in IoT to be addressed are authentication, confidentiality, privacy, and access control. As per [7], prime factors behind the vulnerabilities of IoT can be listed as below:

1. Lack of privacy due to the limited computation capability of the CPU in IoT devices
2. Usage of battery (limited power source) leading to low power consumption for security algorithms
3. Restrictions on having low cost implementation of security algorithms for increasing the number of devices covered
4. Heterogeneous nature of the network which makes it difficult for a uniform solution

For better learning in machine learning methods,distribution bias aware collaborative models like [20] can be deployed in IoT devices.

3 ATTACKS

The security requirements for the the information system is reflected in the requirement of IoT system too. As a matter of fact, the primary intention of the attacker is compromising the privacy, integrity, availability and confidentiality of the system. To defend against this, the prime objective of an IoT ecosystem is to provide privacy, confidentiality, integrity, and availability as given in Fig. 1 below.

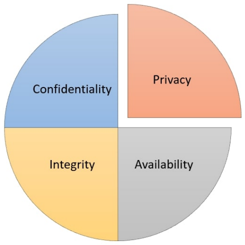

Fig. 1. Prime security concerns in IoT

Attacks in the IoT networks can be classified into two main categories:

1. Passive attacks: Attackers can go undetected since they only observe the information over the network without altering it.
2. Active attacks: Attackers affects the functionality and operation of the system by modifying the information over the network or by directly intervening in the system.

Several notable common passive attacks in IoT networks include eavesdropping, traffic analysis types, sniffing etc. On the other hand, most common active attacks include jamming, flooding, node destruction, node malfunctioning, node outage, man in the middle attacks, targeted code injection, hijacking, denial of service (DoS), and distributed denial of service (DDoS) [8], replay attacks, hole attacks, and Sybil types. On top of that, malware can be elegantly hidden in the large amount of IoT data since there is less scope for verification and authentication. [4] provides a study on the

classification of IoT attacks and their countermeasures. [12] focuses on the network security matters in the smart home, health care and transportation domains. They also construct a taxonomy of security attacks within IoT networks to assist IoT developers for better awareness of the risk of security flaws hoping that better protections to be incorporated. [2] is the most recent survey on the detailed review of security attacks (from active and passive attack point of view) towards WSNs and IoT. They also provides with the techniques for preventing, detecting, and even mitigating those attacks.
Several active attacks that can be detected with IDS are briefly classified here:

3.1 Attacks on OSI Layers

3.1.1 Physical Layer. Jamming, Node tempering, RF interference, micro probing, social engineering, physical damage

3.1.2 MAC Layer. Collusion, exhaustion, de-synchronization, link layer flooding or jamming, spoofing / ARP-spoofing, unfairness, denial/ deprivation of sleep

3.1.3 Network Layer. Hole attacks (blackhole, grayhole, sinkhole, wormhole), node manipulation (replication, destroy, unauthorised access etc), routing attacks (Spoofing, altering, replication or selective forwarding), traffic analysis attacks, Hello flooding, RPL exploit, Sybil attacks, Acknowledgement spoofing, DoS, MITM

3.1.4 Transport Layer. Flooding, de-synchronization, MQTT exploit, session hijacking in TCP, SYN-flooding

3.1.5 Application Layer. False data injection, data aggregation distortion, selective message forwarding, clock skewing, path based DoS, Re-programming, sensor overwhelming, CoAP exploit

3.2 Attacks on Software

Virus, Trojan horse, malware, spyware, malicious node adware, malicious scripts, worms, logic bombs, DoS, reverse engineering

3.3 Attacks on Cryptosystem

3.3.1 Attacks on Side Channels. Timing analysis, power analysis, fault analysis, electromagnetic analysis

3.3.2 Cryptanalysis Attacks . Ciphertext-only attack, Known-Plaintext attack, Chosen-Plaintext attack,

3.3.3 Attacks using a man-in-the-middle (MITM).

3.4 Attacks on Hardware

3.4.1 IoT devices and peripherals. Brute force, Buffer overflow, Rolling code attacks, BlueBorne attack, Sybil attack

3.4.2 Gateways and internal network. Injection attack, MITM, DNS poisoning, Replay attacks, Wormhole

3.4.3 Cloud servers and control devices. SQL-injection, DDoS, weak authentication, malicious applications, Back doors and exploits

4 ALGORITHMS

In this section, we present a comprehensive survey of a variety of algorithms and models that have been applied in recent research for Intrusion Detection Systems (IDS) in IoT. Each algorithm is briefly discussed, with an overview of its functionality, strengths, and weaknesses, and its application in the context of IDS for IoT.

The surveyed algorithms are categorized into three main groups: Traditional Machine Learning Algorithms, Deep Learning Algorithms, and Other Algorithms.

4.1 Traditional Machine Learning Algorithms

These algorithms are widely used due to their interpretability and effectiveness in handling high-dimensional data.

4.1.1 Support Vector Machines (SVMs). Support Vector Machines (SVMs) are supervised learning models used for classification and regression analysis. In the context of IDS, SVMs have been widely used due to their ability to handle high-dimensional data and their robustness against overfitting. According to paper[3], SVMs have been used with Particle Swarm Optimization (PSO) for feature selection, which has proven effective in reducing detection time while maintaining high detection accuracy.

4.1.2 Decision Trees (DTs). Decision Trees (DTs) are a type of supervised learning algorithm that are mostly used for classification problems. They are highly interpretable and can handle both categorical and numerical data. However, they can easily overfit or underfit data, making them sensitive to the dataset. In paper[3], DTs were used in conjunction with PSO, resulting in reduced detection time and high detection accuracy.

4.1.3 K-Nearest Neighbors (K-NN). K-Nearest Neighbors (K-NN) is a type of instance-based learning algorithm. It stores all instances correspondingly and classifies new instances based on a similarity measure. K-NN has been used in intrusion detection systems due to its simplicity and effectiveness. In paper [1], K-NN was used for anomaly detection in IoT devices.

4.1.4 Random Forest. Random Forest is a learning method that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees for classification tasks, or mean prediction of the individual trees for regression tasks. In the context of IDS, as discussed in paper [5], Random Forest is used for classification tasks, contributing to the high accuracy of the proposed model.

4.1.5 Naive Bayes. Naive Bayes is a probabilistic machine learning algorithm that is widely used for classification problems. It is based on Bayes' theorem, which assumes that the features used in classification are conditionally independent of each other. Naive Bayes is computationally efficient and requires a small amount of training data, making it useful for real-time classification tasks. In the context of IDS, Naive Bayes has been used for network intrusion detection, with good results reported in paper [10]. However, its assumption of feature independence may not hold true in some real-world scenarios, leading to lower accuracy compared to other algorithms.

4.2 Deep Learning Algorithms

Deep Learning Algorithms are a subset of machine learning algorithms that have multiple layers between the input and output layers. They can model complex non-linear relationships and are highly scalable.

4.2.1 Deep Neural Networks (DNNs). Deep Neural Networks (DNNs) are a subset of neural networks with multiple layers between the input and output layers. DNNs can model complex non-linear relationships and are highly scalable. However, DNNs require a large amount of data to train effectively, and they are not easily interpretable. In paper [9], DNNs were used for intrusion detection in the Controller Area Network (CAN) bus. Despite their simplicity, DNNs lack the capacity to learn from prior input and training iterations, which is a limitation in their use.

4.2.2 Convolutional Neural Networks (CNNs). Convolutional Neural Networks (CNNs) are a class of deep learning algorithms predominantly used in image processing tasks. CNNs are designed to automatically and adaptively learn spatial hier

archies of features from tasks with input data in 2 or 3 dimensions. In the context of IDS as discussed in paper [9], CNNs require a conversion of CAN frames into images to be applicable. Despite this, their high specialization for image processing can be leveraged effectively in certain intrusion detection scenarios.

4.2.3 Long Short-term Memory (LSTM) networks. Long Short-term Memory (LSTM) networks are a type of Recurrent Neural Network (RNN) capable of learning long-term dependencies. This makes them highly suitable for time-series data, as they can understand context and remember or forget information over long sequences. In the IDS context discussed in paper [9], LSTMs can model temporal dependencies but require a separate LSTM for each arbitration ID in the CAN bus, which could be computationally expensive.

4.2.4 Generative Adversarial Networks (GANs). Generative Adversarial Networks (GANs) are a class of machine learning frameworks designed to generate new, synthetic instances of data that can pass for real instances. They are composed of two components, a generator, and a discriminator, which are trained simultaneously. The generator creates synthetic instances to fool the discriminator, while the discriminator gets better at distinguishing real instances from the fakes. In the context of IDS, as discussed in papers [9] and [5], GANs can help generate abnormal data for training purposes, enhancing the robustness of the model and its ability to handle unbalanced datasets.

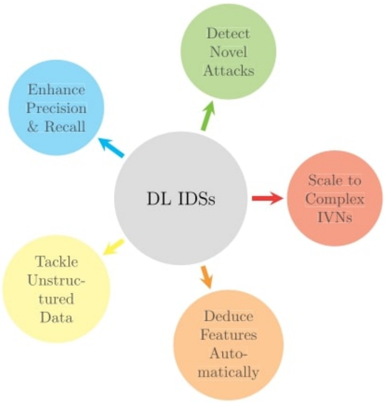

Fig. 2. Advantages of Deep Learning Intrusion Detection Systems in Automotive Applications [9]

4.3 Other Algorithms

This category includes algorithms that do not strictly fall into the traditional machine learning or deep learning categories.

4.3.1 Deep Transfer Learning (DTL). Deep Transfer Learning (DTL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. In the context of IDS as discussed in paper [9], DTL can be used to reduce training time and data requirements by leveraging pre-trained models.

4.3.2 Attention & Transformer models. Attention and Transformer models are a form of deep learning model architecture that focuses on the mechanism of attention, i.e., understanding where to direct focus when processing data. In the context of IDS discussed in paper [9], these models can emphasize informative data and de-emphasize uninformative data, making them potentially effective in intrusion detection tasks.

4.3.3 Denoising Autoencoder (DAE). A Denoising Autoencoder (DAE) is a type of autoencoder, which is a neural network used for learning efficient codings of input data. Specifically, a DAE takes a partially corrupted input and is trained to recover the original undistorted input. In the context of IDS, as discussed in paper [5], a DAE is used for data feature extraction and representation, improving the performance of the proposed model.

4.3.4 Federated Deep Reinforcement Learning. Deep Reinforcement Learning (DRL) combines deep learning and reinforcement learning principles to create efficient algorithms that can be applied to complex learning tasks. Federated learning, on the other hand, is a machine learning approach that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them. As discussed in paper [13], a Federated Deep Reinforcement Learning approach is proposed for traffic monitoring in SDN-based IoT networks, showing promising results.

4.4 Energy Efficient Anomaly Detection Using Game Theory

In some recent studies, game theory has been applied to enhance the efficiency of IDS in resource-constrained IoT devices. One such approach is presented in a study by Sedjelmaci et al. [18], which proposes an anomaly detection system that balances detection accuracy and energy consumption.

The proposed approach combines the strengths of anomaly detection and signature detection, resulting in a high detection rate and a low false positive rate. Game theory, specifically the concept of Nash equilibrium, is utilized to optimize the activation of anomaly detection. According to the study, anomaly detection is only activated when a new attack signature is anticipated to occur. This strategy aims to minimize the energy consumption that typically arises when anomaly detection is continuously active, particularly in low-resource IoT devices.

Simulation results from the study indicate that this approach leads to a reduction in energy consumption while maintaining high detection accuracy (i.e., high detection and low false positive rates). This work provides valuable insights into how game-theoretical models can contribute to enhancing the efficiency of IDS in IoT networks, particularly in scenarios with resource constraints.

4.5 Summary of Algorithms and Techniques

This survey presents an overview of various algorithms and techniques utilized in the design of IDS for IoT systems. Traditional machine learning approaches such as Support Vector Machines, Decision Trees, K-Nearest Neighbors, and Naive Bayes, offer high interpretability and can handle high-dimensional data but may sometimes struggle with overfitting or underfitting, and data imbalance.

Deep learning algorithms such as Deep Neural Networks, Convolutional Neural Networks, and Recurrent Neural Networks, including Long Short-Term Memory networks, provide powerful tools for modeling complex patterns and relationships in data, offering advantages in scenarios where data is abundant and computational resources are ample. However, these algorithms are typically resource-intensive and may be challenging to interpret.

The use of Generative Adversarial Networks for synthetic data generation, Deep Transfer Learning for leveraging pre-trained models, and Denoising Autoencoders for efficient feature extraction and representation have also been explored. Additionally, the integration of Attention and Transformer models into IDS has shown promise in emphasizing informative data and de-emphasizing uninformative data.

Game theory, specifically the concept of Nash equilibrium, has been utilized in IDS for IoT to optimize anomaly detection activation, balancing the trade-off between detection accuracy and energy consumption in low-resource IoT devices.

All these techniques, when appropriately applied, contribute to enhancing the robustness and efficiency of intrusion detection in IoT. However, it's important to acknowledge that each technique has its own strengths and limitations and may be more suitable for certain scenarios over others. Therefore, the choice of algorithm or technique should be made judiciously, considering the specific requirements and constraints of the IoT environment.

We continue to observe research advancements in this field, with ongoing efforts towards developing more efficient, scalable, and robust IDS for IoT. It is anticipated that future studies will continue to explore the use of novel algorithms, hybrid approaches, and optimization techniques, in addition to addressing challenges related to data scarcity, computational constraints, real-time processing, and privacy preservation.

Table 1 provides a brierf comparison of these algorithms.

Algorithm	Strengths	Challenges
SVM	Robust against overfitting, handles high-dimensional data	Requires careful parameter tuning
Decision Tree	Highly interpretable, can handle categorical and numerical data	Can easily overfit or underfit data
Random Forest	Ensemble method, reduces overfitting, handles imbalanced data	Computationally expensive with large datasets
Naive Bayes	Simple, efficient, handles categorical and continuous data	Makes strong independence assumptions
DNN	Can model complex non-linear relationships	Require large amount of data, not easily interpretable
CNN	Effective for pattern recognition in network traffic	Require careful design of input feature representation
LSTM	Can learn long-term dependencies, suitable for time-series data	Computationally expensive for large datasets
GAN	Can generate synthetic instances for training, enhancing model robustness	May experience unstable performance
DTL	Can leverage pre-trained models to reduce training time	Requires large dataset for the pre-training phase
Attention & Transformer models	Can emphasize informative data, de-emphasize uninformative data	Require large amount of data and computational resources
DAE	Good for feature extraction and data representation	Requires careful tuning, can be computationally expensive

Table 1. Comparison of IDS Algorithms for IoT, Strenghts and Challengesik v

5 METHODOLOGY

This section presents an overview of the methodology used for employing machine learning algorithms in the development of Intrusion Detection Systems (IDS) for IoT. The methodology consists of several steps, including defining the problem and system requirements, data collection and preprocessing, feature extraction, model selection, training, evaluation, refinement, deployment, and performance monitoring. By following these steps, researchers and practitioners can effectively develop and deploy IDS solutions tailored to the unique characteristics and requirements of IoT systems. In the following subsections, we will discuss each of these steps in more detail and provide insights into their importance in the context of IDS for IoT.

5.1 Defining the Problem and System Requirements

The first step in the methodology is to clearly define the problem to be solved and the specific requirements of the IoT system. This involves identifying the types of devices, communication protocols, and network architectures used in the IoT environment. Additionally, it is crucial to determine the types of attacks or intrusions the IDS should be capable of detecting, such as Distributed Denial of Service (DDoS) attacks, malware infections, or unauthorized access.

Understanding the system requirements will help guide the selection of data sources, feature extraction techniques, and machine learning models that are most suitable for the given IoT scenario. Moreover, it is essential to define the desired performance metrics for the IDS, such as detection accuracy, false alarm rate, and response time, as these will serve as benchmarks for evaluating and refining the system throughout the development process.

5.2 Data Collection, Preprocessing, and Feature Extraction

Once the problem and system requirements have been defined, the next steps involve collecting and processing the data, and extracting relevant features. These steps are crucial for preparing the data for machine learning models and ensuring their effectiveness in detecting intrusions in IoT environments.

1. **Data Collection:** Collect relevant data from the IoT devices and networks to be monitored. This data can include network traffic, system logs, and IoT-specific data, such as sensor readings or device status information. To ensure diverse and representative data, it is essential to collect data from various sources and under different network conditions, including normal operation and simulated attack scenarios.
2. **Preprocessing:** Clean and preprocess the collected data to eliminate inconsistencies, noise, and missing values. This step includes:
   - Removing irrelevant or redundant data
   - Handling missing or incomplete data
   - Converting data into appropriate formats, such as numerical or categorical representations
   - Normalizing or scaling the data to ensure that all features have comparable ranges and distributions
   - Handling imbalanced data by applying techniques like oversampling, undersampling, or generating synthetic instances using methods such as SMOTE (Synthetic Minority Over-sampling Technique) which has been used in paper [5]
3. **Feature Extraction:** Extract relevant features from the preprocessed data that are useful for detecting intrusions and attacks. This step involves:
   - Identifying important attributes or characteristics of the data, such as packet size, duration, and frequency, that can help discriminate between normal and malicious activities
   - Employing statistical or machine learning techniques, such as Principal Component Analysis (PCA) or autoencoders, to reduce dimensionality and extract meaningful features
   - Evaluating the relevance and importance of the extracted features using techniques like feature ranking or recursive feature elimination

For instance, Paper [16] employed PCA for dimensionality reduction, reducing computational costs and enhancing the algorithm's efficiency. Similarly, Paper [11] used the Fast Fourier Transform (FFT) to extract features from traffic data.

5.3 Model Selection, Training, and Refinement

After preprocessing the data and extracting relevant features, the next steps involve selecting an appropriate machine learning model, training the model using the extracted features, and refining the model based on its performance. These steps ensure that the IDS is capable of effectively detecting intrusions in the IoT environment.

(1) **Model Selection:** Choose a suitable machine learning model for the given problem and data, taking into account the characteristics of the IoT system and the desired performance metrics. Some common models used for intrusion detection include support vector machines, decision trees, deep learning, and ensemble methods. Factors to consider when selecting a model include:
   - The nature of the data (e.g., continuous, categorical, time-series)
   - The size and dimensionality of the dataset
   - The complexity of the relationships between features and classes
   - The desired trade-off between model interpretability and performance

(2) **Model Training:** Train the selected model using the extracted features and labeled data instances. This step typically involves splitting the dataset into training and validation sets, with the training set used to learn the model's parameters and the validation set used to tune the model's hyperparameters and prevent overfitting. Some important aspects of model training include:
   - Regularization techniques to prevent overfitting (e.g., L1 or L2 regularization, dropout)
   - Parameter tuning using methods such as grid search, random search, or Bayesian optimization
   - Leveraging pre-trained models or transfer learning to reduce training time and improve performance, especially when the available data is limited
   - Employing techniques such as cross-validation to ensure robust model evaluation and selection

   For instance, paper [6] uses a 70-30 split for training and testing, respectively, while Paper [19] uses a 60-20-20 split for training, validation, and testing.

(3) **Model Refinement:** Refine the model and its parameters based on the evaluation results, with the goal of improving its performance on the desired metrics (e.g., detection accuracy, false alarm rate, response time). This iterative process may involve adjusting the model architecture, feature set, or training parameters, and can be guided by techniques such as:
   - Learning curves to diagnose issues with model performance (e.g., high bias or high variance)
   - Feature importance analysis to identify potential improvements in feature selection or extraction
   - Sensitivity analysis to understand the impact of individual hyperparameters on model performance

   For instance, Paper [11] adjusts the hyperparameters of the SVM model to enhance its detection capability. Similarly, Paper [3] uses a grid search method to find optimal parameters for the CNN model.

Also, The use of pre-trained models, which have already been trained on large datasets, is also prevalent. These models can be fine-tuned for the specific task of intrusion detection. For example, Paper [9] utilizes a pre-trained LSTM model, which is then fine-tuned to detect intrusions in IoT networks.

5.4 Model Deployment and Performance Monitoring

Once the machine learning model has been trained and refined, it is ready to be deployed in the IoT system for real-time monitoring and intrusion detection. This section discusses the deployment of the IDS and the continuous monitoring of its performance to ensure its effectiveness and adaptability in the face of evolving threats.

(1) **Model Deployment:** Deploy the trained and refined model in the IoT environment to monitor network traffic, system logs, and other relevant data for potential intrusions. This may involve integrating the model into existing IoT devices or gateways, or deploying it on edge or cloud servers that can process the data and communicate with the IoT devices. Some key considerations for model deployment include:
   - Ensuring efficient and scalable processing of the data, especially in large-scale IoT systems
   - Addressing potential security and privacy concerns related to the transmission and processing of sensitive data
   - Providing mechanisms for updating the model and its parameters as new data and threats emerge

(2) **Performance Monitoring:** Continuously monitor the performance of the deployed IDS, assessing its ability to detect intrusions accurately and with minimal false alarms. Performance monitoring enables the identification of potential issues and areas for improvement, such as the need for additional training data or updates to the model's parameters. Techniques and metrics for monitoring performance include:
   - Tracking detection accuracy, precision, recall, F1-score, and Area Under the ROC Curve (AUC)
   - Monitoring the false alarm rate, response time, and computational resources consumed by the IDS
   - Comparing the performance of the IDS to other models or benchmarks in the literature or industry
   - Analyzing the types of attacks or intrusions that the IDS may be missing or misclassifying to inform future refinements

5.5 Performance Comparison

To understand the effectiveness of the models, their performance is often compared with other models or methods. This comparison helps to quantify the improvements achieved by the proposed models.

- **Comparing with Other Models:** Many of the reviewed papers compare their proposed models with existing models to highlight their improvements. For instance, Paper [3] compares the performance of their proposed CNN model with a traditional SVM model, showing significant improvements.
- **Improvement over Other Methods:** The papers also show improvements over other methods. For instance, Paper [19] shows that their ensemble model provides an improvement in detection accuracy over individual models.

Developing an effective Intrusion Detection System (IDS) for IoT environments is a challenging task that requires a systematic approach. By following the steps outlined in this methodology, researchers and practitioners can ensure that they address the key aspects of data collection, preprocessing, feature extraction, model selection, training, refinement, deployment, and performance monitoring. This comprehensive process helps to create a robust, accurate, and adaptable IDS that can protect IoT systems from a wide range of threats and adapt to the evolving landscape of cybersecurity risks.

As new machine learning algorithms and techniques emerge, it is essential to continually re-evaluate and refine the methodologies used in IDS development. By staying up-to-date with the latest research and best practices, researchers and practitioners can contribute to the ongoing improvement of IoT security and the development of more effective and resilient IDS solutions.

6 APPLICATIONS

Emerging machine learning algorithms have shown significant promise in addressing the challenges associated with developing effective Intrusion Detection Systems (IDS) for IoT environments. In this section, we discuss the applications of these algorithms in various aspects of IDS for IoT, including privacy-preserving data analysis, intrusion detection, resource allocation, and traffic monitoring.

6.1 Privacy-Preserving Data Analysis

One of the primary applications of federated learning and GANs in IoT IDS is privacy-preserving data analysis. The P-FedGAN framework proposed in [19] integrates differential privacy (DP) with GANs in a federated learning setting to generate realistic, high-quality synthetic data while preserving the privacy of the original data. This approach is particularly suitable for IoT environments, where sensitive data from multiple sources need to be analyzed for intrusion detection while maintaining user privacy. Differential privacy provides a mathematical guarantee of privacy, making this approach a promising solution for privacy-preserving IDS in IoT applications.

6.2 Intrusion Detection

Effective security solutions, including intrusion detection systems, are vital in IoT due to the exponential increase in the number of connected devices and potential security vulnerabilities. Machine learning algorithms such as SVM, Decision Trees, DNNs, CNNs, LSTMs, and GANs have been applied in various studies to enhance the effectiveness of intrusion detection in IoT systems. In [5], a hybrid model combining machine learning and deep learning was proposed for intrusion detection in a 5G IoT environment. The model, which uses a GAN network for data optimization and imbalance processing, achieved high detection accuracies, demonstrating the potential of these algorithms in enhancing the security of IoT systems.

6.3 Resource Allocation

Effective resource allocation is crucial in IoT networks to ensure optimal system performance, particularly in the context of IDS. Machine learning techniques like federated learning and DRL can be instrumental in this regard. For instance, in [17], a federated multi-agent actor-critic algorithm was proposed for joint radio and computational resource allocation in multi-cell Mobile Edge Computing (MEC) networks. The proposed model demonstrated superior performance in terms of the system's energy efficiency, showcasing the potential of federated learning and DRL in enhancing the efficiency of IoT IDS.

6.4 Traffic Monitoring

Traffic monitoring is a critical task for intrusion detection in IoT networks, as it enables the identification of anomalies and potential attacks. Machine learning techniques like federated learning and DRL can play a significant role in this regard. The DeepMonitor framework proposed in [13] utilized federated deep reinforcement learning to monitor traffic efficiently in SDN-based IoT networks. The proposed framework achieved better performance than traditional methods, demonstrating the effectiveness of federated learning and DRL in traffic monitoring tasks in IoT networks. Table 2 summarizes the applications of emerging ML algorithms in IoT IDS, highlighting the techniques used and their corresponding references.

Application	Techniques Used	Reference
Privacy-Preserving Data Analysis	Federated Learning, GANs	[19]
Intrusion Detection	Federated Learning, GANs	[5], [13]
Resource Allocation	Federated Learning, DRL	[17]
Traffic Monitoring	Federated Learning, DRL	[13]

Table 2. Summary of Applications of Emerging ML Algorithms in IDS for IoT in the reviewed papers

In conclusion, emerging machine learning algorithms show promise in various IoT IDS applications, making these techniques instrumental in the future of IoT systems. Each approach's inherent advantages, such as privacy preservation in federated learning, realistic synthetic data generation in GANs, and efficient decision-making in DRL, can be leveraged to address the unique challenges in IoT intrusion detection. As IoT networks continue to grow and evolve, these techniques' role will undoubtedly become increasingly important.

Despite the demonstrated successes, numerous challenges and research opportunities remain in applying these techniques in IoT IDS. For instance, further research is needed to improve the efficiency and scalability of federated learning algorithms in large-scale IoT networks. Similarly, enhancing the realism of synthetic data generated by GANs without compromising privacy remains a significant challenge. In the context of DRL, developing algorithms capable of handling the complexity and dynamism of IoT networks is a critical research direction. As we move towards an increasingly connected future, tackling these challenges will be paramount in fully realizing the potential of emerging machine learning algorithms in IoT IDS.

7 DEFENSE

Prevention, detection, and mitigation these 3 essential components forms the line of defense against security attacks on IoT networks.

7.1 Prevention

Some effective ways to prevent attacks in IoT networks:

- Securing all devices with strong passwords
- Authorizing devices through unique fingerprint defined by physical state and location
- Applying software defined networking,
- Using antivirus and firewalls
- Regularly updating software, firmware, apps, and operating systems
- Hiding network traffic by using Synthetic Packet Engine (SPE) to generate fake additional traffic data
- Utilizing monitoring tools to detect the intrusion
- Deploying lightweight encryption algorithm to maintain the confidentiality of the data
- Encryption query process algorithm may perform better than encryption in some cases

7.2 Detection

Notable techniques for detection mechanism:

- Identifying the point of contact that is compromised
- Utilizing Intrusion Detection Systems for reacting against ongoing attacks

7.3 Mitigation

Actions taken to mitigate the impact of attacks:

- Dismissing the compromised component in the network
- Disabling the affected device from communicating with others

8 CONCLUSION

IoT technologies consist of a complex network of smart devices that continuously exchanges data over the network. IoT system are ever expanding in our day to day life because of their ubiquitous use in automation. As a result, we must handle its security issues regarding safety critical operation, exposure of sensitive data, and secure remote access with proper attention. Machine learning models with better customization for application scenario can benefit such prevention, and detection of such security attacks. Focusing on different objective function such as cost of implementation, resources available, etc separate machine learning approaches may be incorporated. There is no one for all model that is suitable everywhere for various privacy and security issues in IoT ecosystem. Future works can focus on securing decentralized IoT systems with low cost but efficient security protocols implementation.

REFERENCES

[1] Mansoor Ahmed Bhatti, Rabia Riaz, Sanam Shahla Rizvi, Sana Shokat, Farina Riaz, and Se Jin Kwon. 2020. Outlier detection in indoor localization and Internet of Things (IoT) using machine learning. *Journal of Communications and Networks* 22, 3 (2020), 236–243. https://doi.org/10.1109/JCN.2020.000018

[2] Ismail Butun, Patrik Österberg, and Houbing Song. 2019. Security of the Internet of Things: Vulnerabilities, attacks, and countermeasures. *IEEE Communications Surveys & Tutorials* 22, 1 (2019), 616–644.

[3] Kelton AP Da Costa, João P Papa, Celso O Lisboa, Roberto Munoz, and Victor Hugo C de Albuquerque. 2019. Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks* 151 (2019), 147–157.

[4] Jyoti Deogirikar and Amarsinh Vidhate. 2017. Security attacks in IoT: A survey. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*. IEEE, 32–37.

[5] S. Ding, L. Kou, and T. Wu. 2022. A GAN-Based Intrusion Detection Model for 5G Enabled Future Metaverse. *Mobile Networks and Applications* 27 (2022), 2596–2610. https://doi.org/10.1007/s11036-022-02075-6

[6] Aidin Ferdowsi and Walid Saad. 2019. Generative adversarial networks for distributed intrusion detection in the internet of things. In *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 1–6.

[7] Ali Hameed and Alauddin Alomary. 2019. Security issues in IoT: a survey. In *2019 International conference on innovation and intelligence for informatics, computing, and technologies (3ICT)*. IEEE, 1–5.

[8] Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou, and Jeffrey Voas. 2017. DDoS in the IoT: Mirai and other botnets. *Computer* 50, 7 (2017), 80–84.

[9] Brooke Lampe and Weizhi Meng. 2023. A survey of deep learning-based intrusion detection in automotive applications. *Expert Systems with Applications* 221 (2023), 119771. https://doi.org/10.1016/j.eswa.2023.119771

[10] Tahir Mehmood and Helmi B. Md Rais. 2016. Machine learning algorithms in context of intrusion detection. In *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*. 369–373. https://doi.org/10.1109/ICCOINS.2016.7783243

[11] Viraaji Mothukuri, Prachi Khare, Reza M. Parizi, Seyedamin Pouriyeh, Ali Dehghantanha, and Gautam Srivastava. 2022. Federated-Learning-Based Anomaly Detection for IoT Security Attacks. *IEEE Internet of Things Journal* 9, 4 (2022), 2545–2554. https://doi.org/10.1109/JIOT.2021.3077803

[12] Mukrimah Nawir, Amiza Amir, Naimah Yaakob, and Ong Bi Lynn. 2016. Internet of Things (IoT): Taxonomy of security attacks. In *2016 3rd international conference on electronic design (ICED)*. IEEE, 321–326.

[13] Tri Gia Nguyen, Trung V Phan, Dinh Thai Hoang, Tu N Nguyen, and Chakchai So-In. 2021. Federated deep reinforcement learning for traffic monitoring in SDN-based IoT networks. *IEEE Transactions on Cognitive Communications and Networking* 7, 4 (2021), 1048–1065.

[14] Gowthamaraj Rajendran, R S Ragul Nivash, Purushotham Parthiban Parthy, and S Balamurugan. 2019. Modern security threats in the Internet of Things (IoT): Attacks and Countermeasures. In *2019 International Carnahan Conference on Security Technology (ICCST)*. 1–6. https://doi.org/10.1109/CCST.2019.8888399

[15] José Roldán, Juan Boubeta-Puig, José Luis Martínez, and Guadalupe Ortiz. 2020. Integrating complex event processing and machine learning: An intelligent architecture for detecting IoT security attacks. *Expert Systems with Applications* 149 (2020), 113251.

[16] Mahdis Saharkhizan, Amin Azmoodeh, Ali Dehghantanha, Kim-Kwang Raymond Choo, and Reza M Parizi. 2020. An ensemble of deep recurrent neural networks for detecting IoT cyber attacks using network traffic. *IEEE Internet of Things Journal* 7, 9 (2020), 8852–8859.

[17] Leonel Santos, Carlos Rabadao, and Ramiro Gonçalves. 2018. Intrusion detection systems in Internet of Things: A literature review. In *2018 13th Iberian conference on information systems and technologies (CISTI)*. IEEE, 1–7.

[18] Hichem Sedjelmaci, Sidi-Mohammed Senouci, and Mohamad Al-Bahri. 2016. A lightweight anomaly detection technique for low-resource IoT devices: A game-theoretic methodology. *2016 IEEE International Conference on Communications (ICC)* (2016), 1–6.

[19] Hui Wang, Yani Han, Shaojing Yang, Anxiao Song, and Tao Zhang. 2021. Privacy-Preserving Federated Generative Adversarial Network for IoT. In *2021 International Conference on Networking and Network Applications (NaNA)*. IEEE, 75–80.

[20] Xiaokang Zhou, Yiyong Hu, Jiayi Wu, Wei Liang, Jianhua Ma, and Qun Jin. 2022. Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT. *IEEE Transactions on Industrial Informatics* 19, 1 (2022), 570–580.

[21] Tian Zixu, Kushan Sudheera Kalupahana Liyanage, and Mohan Gurusamy. 2020. Generative adversarial network and auto encoder based anomaly detection in distributed IoT networks. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 1–7.

Received 04 May 2023