International Journal of Electrical, Electronics and Computer Engineering **14**(1&2): 51-54(2025)

ISSN No. (Online): 2277-2626

Anomaly Detection in IoT Networks Using Machine Learning

Chahat Sharma* and Muskan Thakur
School of Computer Science Engineering and Technology,
Government College Dharamshala (H.P.), India.

(Corresponding author: Chahat Sharma)*
(Received: 01 March 2025, Accepted: 05 April 2025)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The growth of IoT networks has brought new security risks, especially in detecting unusual activities that may indicate attacks or failures. Traditional security methods struggle with the dynamic nature of IoT environments. This paper explores the use of Machine Learning (ML) techniques, such as Support Vector Machines (SVM), Decision Trees, and Clustering, for detecting anomalies in IoT networks. The focus is on developing automated and scalable systems to improve IoT security and reliability, along with suggestions for overcoming existing challenges.

Keywords: IoT, Anomaly Detection, Supervised learning, Unsupervised learning, SVM, Decision Trees.

INTRODUCTION

The Internet of Things (IoT) is rapidly transforming the way devices communicate and interact with each other. IoT networks consist of interconnected devices such as sensors, cameras, smart appliances, and wearable gadgets, which exchange data over the internet. While IoT has made significant advancements in various sectors like healthcare, transportation, and smart homes, it has also introduced new security concerns. Due to the large number of connected devices and their limited computational capabilities, IoT networks are vulnerable to cyber-attacks and unauthorized access.

One of the most critical challenges in securing IoT networks is the detection of anomalies. Anomalies refer to any unusual behavior or irregular patterns in network traffic that may indicate security breaches, malfunctioning devices, or malicious activities. Traditional security systems, such as rule-based intrusion detection systems (IDS), often struggle to handle the dynamic and diverse nature of IoT environments. To address this issue, Machine Learning (ML) techniques have emerged as a powerful solution. ML algorithms can analyze large volumes of network data, learn normal behavior patterns, and effectively detect deviations or anomalies. This paper focuses on the application of machine learning methods for anomaly detection in IoT networks, aiming to enhance network security and reliability.

RELATED WORK

Anomaly detection in IoT networks has been a growing area of research, with several studies focusing on the application of machine learning techniques to address the unique challenges posed by these systems. The inherent complexity, large scale, and heterogeneity of IoT networks make them vulnerable to a variety of attacks, including data manipulation, device compromise, and Distributed Denial of Service (DDoS) attacks. Traditional Intrusion Detection Systems (IDS) have been commonly used to detect anomalies in networks, but these methods rely heavily on rule-based or signature-based techniques. While rule-based IDS can effectively identify attacks that match known signatures, they are limited in their ability to detect new or unknown attacks, which are a constant threat in evolving IoT networks. Moreover, these systems require frequent manual updates to incorporate new attack patterns, which is not always feasible in dynamic and large-scale IoT environments. To overcome these limitations, machine learning-based approaches have gained prominence for anomaly detection in IoT networks. Unlike traditional methods, machine learning can autonomously learn from data, adapt to new attack patterns, and provide a more dynamic and scalable solution. Machine learning techniques, especially supervised learning models, have shown great potential in detecting known anomalies by training on labeled data. However, their reliance on labeled datasets can be a significant hurdle, as labeled data is often hard to obtain in real-world IoT deployments, where attacks may be rare or novel.

Supervised Learning Approaches: Alrashdi and Alsheikh (2019) proposed a supervised learning model for detecting anomalies in smart home IoT networks TURMusing Support Vector Machines (SVM) and Decision Trees. Their model aimed to identify abnormal behavior in network traffic generated by smart home devices, such as unauthorized access or device malfunctions. The approach demonstrated high accuracy in detection, showing that machine learning algorithms could outperform traditional IDS. However, the challenge of scalability remained prominent. As IoT

networks grow in size, with hundreds or thousands of connected devices, training supervised models with sufficient labeled data becomes increasingly difficult and resource-intensive. The scalability issue arises because these models struggle to process and analyze data from large, diverse IoT networks in real time, an essential requirement for practical deployment.

Deep Learning Models: Similarly, Meidan and Shabtai (2018) introduced a deep learning model that utilized neural networks to detect compromised IoT devices based on their network traffic behavior. This model aimed to learn the normal traffic patterns of IoT devices and identify deviations that could indicate security breaches or malware infections. Deep learning models, particularly those based on neural networks, are well-suited for capturing complex and non-linear relationships in data, making them highly effective for anomaly detection.

However, deep learning models require significant computational power, which is often not available in resource-constrained IoT devices. Furthermore, these models tend to consume more memory and energy, which makes them unsuitable for IoT environments where low power consumption is a critical concern. Although deep learning approaches offer high detection accuracy, the trade-off between accuracy and resource consumption poses a significant challenge.

Ensemble and Hybrid Methods: Doshi and Patil (2018) applied ensemble learning techniques, specifically Random Forest and K-Nearest Neighbor (KNN) algorithms, to detect DDoS attacks in IoT networks. DDoS attacks, which aim to overwhelm devices or networks with traffic, pose a significant threat to the reliability and availability of IoT systems. The study demonstrated that ensemble methods could provide a more robust solution for detecting DDoS attacks due to their ability to combine the predictions of multiple models and improve overall accuracy. Random Forest, for example, works by creating a collection of decision trees and then aggregating their results. This approach improved detection rates, especially in heterogeneous IoT environments.

However, despite the high detection rates, the study highlighted challenges in processing large volumes of real-time data. IoT networks generate vast amounts of data, and processing this data in real time for anomaly detection is a significant challenge, particularly for ensemble models that may require more computational resources. In another study, (Aprilia and Gunawan 2020) employed unsupervised learning techniques, specifically K-Means clustering, to detect unknown attacks in IoT networks. Unlike supervised methods, unsupervised methods do not require labeled data, which makes them highly applicable in scenarios where labeled datasets are scarce or difficult to obtain. K-Means clustering works by grouping similar data points and identifying those that deviate significantly from the clusters as potential anomalies. While K-Means proved useful in detecting novel or unknown attacks, it was prone to a higher rate of false positives. This is because K-Means relies solely on the distance between data points, and outliers that do not conform to the established clusters may be flagged as anomalies, even if they are benign. Reducing false positives remains an ongoing challenge for unsupervised methods.

Ensemble Learning and Imbalanced Datasets: Javaid and Anwar (2016) employed ensemble learning techniques to improve detection accuracy in IoT anomaly detection. Ensemble learning combines multiple base models to generate a more accurate and robust prediction. By using multiple algorithms, such as Decision Trees, SVM, and KNN, Javaid and Anwar (2016) sought to improve the performance of anomaly detection systems in IoT networks. Despite improvements in accuracy, one major issue that their study highlighted was the challenge of dealing with imbalanced datasets. In IoT networks, some types of attacks are far rarer than others, leading to imbalanced data where normal activities vastly outnumber anomalies. This imbalance can make it difficult for models to detect less frequent attacks, resulting in reduced overall performance.

Hybrid Approaches: Lastly, Roy and Ghosh (2021) presented a hybrid model that combined both supervised and unsupervised learning techniques. By integrating supervised methods (such as SVM) with unsupervised clustering techniques (such as K-Means), their approach aimed to improve detection accuracy while minimizing false alarms. The hybrid model was able to leverage the advantages of both types of learning, offering a more comprehensive solution for detecting a wide variety of anomalies in IoT networks. Despite its advantages, the hybrid approach still faced challenges related to model complexity, computational overhead, and real-time processing. The combination of multiple models requires more processing time, which can hinder its application in environments where low-latency detection is critical.

Challenges and Future Directions: Despite significant progress, many challenges persist in anomaly detection for IoT networks. Scalability remains a major issue, as the growing number of IoT devices demands models that can efficiently handle large amounts of data. Additionally, the real-time nature of IoT applications requires models that can provide timely and accurate anomaly detection without introducing significant delays. Reducing false positives and adapting models to dynamic environments are also important areas for improvement. IoT networks are constantly evolving, with devices frequently joining or leaving the network, making it essential for anomaly detection systems to be adaptive and capable of handling such changes.

Table 1: Comparative study of existing techniques.

Author, Publisher, and Year	Technique Used	Objective	Performance Metrics	Dataset	Simulator/Outcomes
Alrashdi and Alsheikh (2019)	Support Vector Machines (SVM), Decision Trees	Detect anomalies in smart home IoT networks, identify unauthorized access or device malfunctions.	High accuracy in detection, outperformed traditional IDS.	Network traffic data from smart home devices.	Demonstrated high accuracy but faced scalability issues.
Meidan and Shabtai (2018)	Deep Learning (Neural Networks)	Detect compromised IoT devices based on network traffic behavior.	High detection accuracy, effective in capturing complex patterns.	Network traffic data from IoT devices.	High accuracy but requires significant computational power, unsuitable for low-power IoT devices.
Doshi and Patil (2018)	Ensemble Learning (Random Forest, K-Nearest Neighbor - KNN)	Detect DDoS attacks in IoT networks.	Improved detection rates, robust in heterogeneous IoT environments.	Network traffic data from IoT devices.	Improved accuracy but faced challenges in real-time data processing.
Aprilia and Gunawan (2020)	Unsupervised Learning (K-Means Clustering)	Detect unknown attacks in IoT networks.	Useful for detecting novel attacks, but prone to false positives.	Network traffic data from IoT devices.	No need for labeled data, but higher false positive rates.
Javaid and Anwar (2016)	Ensemble Learning (Decision Trees, SVM, KNN)	Improve detection accuracy in IoT anomaly detection.	Improved accuracy, but struggled with imbalanced datasets.	Network traffic data from IoT devices.	Improved accuracy but faced challenges with imbalanced datasets.
Roy and Ghosh (2021)	Hybrid Model (Supervised: SVM, Unsupervised: K-Means Clustering)	Improve detection accuracy while minimizing false alarms.	Combined advantages of supervised and unsupervised learning, improved accuracy.	Network traffic data from IoT devices.	Improved accuracy but faced challenges in model complexity and real-time processing.

RESEARCH GAP

Even though there has been progress in using machine learning for anomaly detection in IoT networks, there are still several challenges to address. One of the main issues is scalability. As IoT networks grow, with more and more devices being added, existing machine learning models often struggle to handle the large amounts of data generated in real-time. These models also tend to be slow and require a lot of computing power, which can be a problem in fast-changing IoT environments.

Another challenge is the availability of labeled -data. Training machine learning models often requires labeled examples of both normal and abnormal data. However, it is difficult and time-consuming to gather enough labeled data, especially for rare or new types of attacks. This makes it harder to train accurate models, and new approaches are needed that don't rely so much on labeled data.

Real-time anomaly detection is another important gap. In many IoT systems, such as in healthcare or industrial settings, it is crucial to detect threats quickly. Current models, especially deep learning models, often take too long to process data, which makes them unsuitable for time-sensitive tasks. Moreover, IoT networks are made up of many different types of devices that produce different kinds of data. Existing models often struggle to deal with this variety and are not very good at combining information from these different sources.

Lastly, many models are not adaptable to changes in the IoT network. Devices are constantly being added or removed, and the data generated by the network can change over time. This requires models that can adjust to these changes and continue to provide accurate detection. More research is needed to develop models that can handle these challenges effectively.

FINDING SUGGESTIONS

From the research and analysis of existing work on anomaly detection in IoT networks using machine learning, a few important findings have emerged. One of the key observations is that while machine learning models, particularly supervised methods, are effective in detecting known anomalies, they face challenges when applied to large-scale IoT networks. These models struggle to handle the huge volume of data that IoT devices generate. The complexity and real-time nature of IoT networks make it difficult for these models to work efficiently. Another major finding is that most current anomaly detection systems rely heavily on labeled data. However, labeling data for every possible attack is difficult, especially for new or rare threats. This makes it hard for the models to detect new, unknown attacks.

Another challenge is the variety of devices in IoT networks. These networks include devices that produce different kinds of data, such as sensor data, logs, and network traffic. Each type of data needs to be processed

and analyzed differently, which makes it difficult to build a single machine learning model that can handle all types of data effectively. Furthermore, many machine learning models, especially those based on deep learning, require a lot of computational power. This makes them unsuitable for IoT devices, which are often limited in resources such as processing power and energy. Based on these findings, several suggestions can help improve anomaly detection in IoT networks. First, there is a need to develop models that can scale to handle large amounts of data. These models should be able to process data in real-time without using too many resources. To address the challenge of labeled data, researchers could explore unsupervised or semi-supervised learning methods, which do not require large amounts of labeled data. These methods could help the models detect new or unknown attacks that were not previously seen.

Another suggestion is to create hybrid models that combine different machine learning techniques. By combining supervised, unsupervised, and even deep learning methods, these hybrid models could improve detection accuracy while reducing the number of false positives. This would help ensure that the system can identify threats without overwhelming users with false alarms. Additionally, techniques like edge computing or federated learning could be explored to help with the computational limitations of IoT devices. These techniques allow data to be processed closer to the source, reducing the need to send large amounts of data to central servers and improving real-time detection.

Lastly, improving the explainability of machine learning models is another important area. In critical IoT applications like healthcare or industrial systems, it is essential that the decisions made by the anomaly detection system can be understood. This transparency helps build trust in the system and ensures that users can act quickly if an anomaly is detected.

CONCLUSIONS

This research explored how machine learning can be used to detect unusual activities in IoT networks. We found that while machine learning works well for finding known attacks, it has some challenges when dealing with large, changing IoT networks. These challenges include handling lots of data in real-time, needing labeled data, and managing the different types of data from various IoT devices. Also, many models need too much computing power, which is a problem for devices with limited resources.

However, machine learning still has great potential for improving security in IoT networks. To overcome these problems, future work should focus on methods that don't need a lot of labeled data, like unsupervised learning. Combining different techniques and using lighter, more efficient models can also help make anomaly detection faster and more reliable. Additionally, using edge computing can improve detection speed without putting too much strain on IoT devices. In conclusion, while there are challenges, machine learning can greatly improve the security of IoT networks. With more research, we can create better systems to protect IoT networks from different threats.

REFERENCES

Alrashdi, A., & Alsheikh, M. (2019). IoT Anomaly Detection using Support Vector Machines and Decision Trees. *International Journal of Computer Science and Network Security.*

Aprilia, S., & Gunawan, M. (2020). K-Means Clustering for Unsupervised Attack Detection in IoT Networks. *Proceedings of the International Conference on IoT Security.*

Doshi, S., & Patil, P. (2018). Using Random Forest and K-Nearest Neighbor for DDoS Attack Detection in IoT Networks. Journal of Network Security.

Javaid, M., & Anwar, A. (2016). Ensemble Learning Techniques for Improved Anomaly Detection in IoT. *International Journal of Artificial Intelligence.*

Meidan, Y., & Shabtai, A. (2018). Deep Learning Models for IoT Device Anomaly Detection. Journal of Machine Learning and Data Mining.

Roy, S., & Ghosh, A. (2021). Hybrid Approach for IoT Anomaly Detection: Combining Supervised and Unsupervised Methods. *Journal of IoT Security.*