International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

Enhancing Cyber Security through Machine Learning-Based Anomaly Detection in IoT Networks

Dr. Shreyas J[1], Dr. Sudhakar K[2], Dr. I. Bhuvaneshwarri[3], Thamaraiselvan B[4], Lakshmi.M[5], Dr Pooja Nayak S[6]
[1]Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Karnataka 576104
shreyas.j@manipal.edu
[2]B.N.M. Institute of Technology, Bengaluru
ksudhakar.cs@bnmit.in
[3]Government College of Engineering, Erode
ibw@gcee.ac.in
[4]K S Rangasamy Institute of Technology, Tiruchengode
thamarai19799@gmail.com
[5]Nitte Meenakshi Institute of Technology, Bangalore
lakshmi.m@nmit.ac.in
[6]Dayananda Sagar Academy of Technology and Management
Pooja-ise@dsatm.edu.in

Abstract: The rapid proliferation of IOT (Internet of Things) networks has brought transformative benefits to industries and everyday life. However, it has also introduced unprecedented cyber security challenges, necessitating advanced techniques for anomaly detection. This research focuses on enhancing cyber security through the application of machine learning-based anomaly detection methods, specifically One-Class Support Vector Machine (SVM) and Isolation Forest, in the context of IOT networks. While Isolation Forest effectively isolates anomalies by building isolation trees, One-Class SVM models the normal data distribution, effectively separating anomalies. To provide a strong security framework for IoT networks, we suggest a comprehensive strategy that combines both algorithms. Our method enables the detection of anomalies in real-time IOT data streams, facilitating prompt responses to new threats. Data collection, preprocessing, and model training are key components. This study helps protect IOT ecosystems and maintain data integrity and privacy in an increasingly connected world by utilizing the benefits of One-Class SVM and Isolation Forest.

Keywords: Machine Learning, Anomaly Detection ,Cyber Security, One Class SVM, Isolation Forest , Network Security.

I. Introduction:

The internet of things, or IoT, is a system of interconnected devices that connect to one another and share data with the cloud and other IoT devices. IoT simply describes how common objects and devices can gather, exchange, and process data without the need for human intervention by being internet-connected. Various opportunities have been made possible by this interconnection, including the development of smart homes and cities as well as industrial automation and improvements in medicine. However, as IoT spreads further, it has also revealed a number of important security issues that require our attention. IoT includes a huge ecosystem of gadgets, from tiny sensors and actuators to substantial industrial machinery and bright appliances. These devices have sensors, communication components, and computing abilities that enable them to collect information, interact with other devices, and make independent choices. This information can be used for a variety of things, including boosting convenience, increasing efficiency, and enabling insights based on data.

IOT Security challenges:

The rapid proliferation of IoT devices, which expands the attack surface, the resource constraints of many IoT devices, the lack of adequate authentication mechanisms that expose devices to unauthorized access, the collection and transmission of sensitive data, concerns over data privacy, the incompatibility of devices from various manufacturers, and communication problems are just a few of the challenges that IoT cyber security must overcome.

Device Proliferation: Cybercriminals have a huge attack surface thanks to the IoT devices quick proliferation. Each of these gadgets, which can be anything from straightforward sensors to intricate industrial machinery, represents a potential point of entry into a network. The task of securing the numerous devices is onerous.

Limited Resources: Many IoT devices are resource-constrained, which means they have limited energy, memory, and processing speed. Strong security measures are difficult to implement due to this restriction. Devices may

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

become more susceptible to resource-intensive attacks such as brute force and denial-of-service (DoS) attacks as a result.

Firmware Updates : Fire ware updates are necessary on a regular basis to fix bugs and enhance security. Many IoT devices do not, however, have the ability to automatically and seamlessly update. Devices are then susceptible to known security flaws because users might not be aware of the need for updates or might find it difficult to carry them out.

Physical Security: Access to IoT devices physically can be used maliciously. Attackers who successfully gain physical access to a device can tamper with it or steal it, potentially exposing confidential information or jeopardizing network security. It is essential to secure IoT devices in physical locations.

Vulnerabilities in the supply chain: The complexity of the supply chains for IoT devices creates potential points of compromise from production to distribution. Device vulnerabilities can be introduced by insecure parts or malicious actors in the supply chain.

Legacy Hardware: Many IoT deployments use hardware from the past that was not created with security in mind. These outdated devices may be unable to support current security protocols or even receive security updates, making them particularly vulnerable.

Regulatory Compliance: It can be difficult for IoT deployments to comply with legal requirements for data security and protection, such as GDPR or HIPAA. Having IoT systems comply with these rules adds another level of complexity.

II. Literature review:

Khatib, Amine & Hamlich, Mohamed & Hamad, Denis. (2021) et.al The author discuss the importance of cyber security in IoT networks in this paper along with how machine learning models can be used to precisely predict attacks. The paper presents several studies that have been conducted using machine learning algorithms for intrusion detection in IoT networks. In the studies, various types of attacks, such as User to Root (U2R) and Remote to Local (R2L) attacks, are detected through a variety of techniques, including dimension reduction, linear discriminant analysis, and binary classification. The SMOTE technique will be addressed in this paper. The paper concludes with a comparison of the various techniques employed and how well they performed in identifying attacks in IoT networks.

the author discussed the difficulties of defending Cyber Physical Systems (CPS) against online threats as well as the use of machine learning techniques like deep learning to automate anomaly detection. The paper also discusses the need for effective security measures and how deep learning is susceptible to hostile attacks. The suggested methodology uses a novel neural network-based technique to retrain the model with adversarial samples in order to mitigate attacks. The findings reveal that using the suggested defense strategy increased the model's robustness by 92.8%. Future research and suggestions for enhancing the security of machine learning-based anomaly detection in CPS environments are provided in conclusion.

Hafsa Benaddi , Mohammed Jouhari , Khalil Ibrahimi , Jalel Ben Othman , El Mehdi Amhoud(2022)et.al Anomaly Detection in Industrial IoT Using Distributional Reinforcement Learning and Generative Adversarial Networks The Industrial Internet of Things (IIoT) anomaly detection method proposed by the author makes use of Generative Adversarial Networks (GANs) and Distributional Reinforcement Learning (DRL). The authors talk about the difficulties with security in the IIoT and contrast their strategy with current practices. Additionally, they offer a thorough justification for their suggested framework and assess its efficiency using a current intrusion detection dataset.

III. Proposed Method:

We use the Isolation Forest algorithm and the One-Class Support Vector Machine (SVM) in this proposed method for anomaly detection in IoT networks with the goal of enhancing cybersecurity. We first gather and preprocess data from IoT devices, making sure it is cleaned and formatted correctly. Then, in order to take advantage of this algorithm's effective partitioning properties and identify anomalies with shorter tree traversal paths, we train an isolation forest model. As a complementary method to identify anomalies based on departures from the learned distribution, we also use a One-Class SVM to model the normal data distribution simultaneously. These two approaches can be combined to create a thorough and reliable anomaly detection system.

Fig 1: One class SVM for anomaly detection

Anomaly detection is a crucial aspect of cybersecurity in IoT (Internet of Things) networks, where identifying

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

unusual behavior or deviations from the norm can be indicative of security threats. One-Class Support Vector Machine (SVM) is a powerful machine learning technique that can contribute significantly to enhancing cybersecurity in IoT networks. Here's how One-Class SVM can be applied for anomaly detection in IoT networks:

1. **Data Collection:**
   In the context of IoT network cyber security, data collection entails the methodical gathering of data from a wide range of connected devices and sensors within the network. This procedure involves gathering a variety of data types, including communication patterns, sensor readings, device status information, and network traffic logs. IoT networks produce enormous amounts of data streams that can give insights into network behavior and potential anomalies, so data collection must be ongoing and real-time to ensure thorough cyber security monitoring. Organizations are able to proactively identify and mitigate security threats and protect the integrity and privacy of their IoT ecosystems thanks to the collected data, which also serves as the fundamental input for subsequent data preprocessing, machine learning model training, and anomaly detection.
2. **Data Preprocessing:**
   In the context of IoT network cyber security, data preprocessing is a crucial step in getting raw data ready for analysis. To make sure that the collected data is of high quality and appropriate for effective analysis, this process entails a number of steps, including data cleaning, transformation, and normalization. The handling of missing values, locating and dealing with outliers, minimizing noise, and converting data into a consistent format are all included in data preprocessing. It also includes feature engineering, a process that can entail extracting pertinent data, choosing crucial features, and reducing dimensionality. These actions all help to increase the accuracy and effectiveness of ensuing machine learning-based anomaly detection models. At the end of the day, data preprocessing ensures that the data is trustworthy, pertinent, and conducive to insightful insights and threat detection.
3. **Model training:**
   Model training, a pivotal phase in machine learning, involves the process of teaching a computational model, often referred to as an algorithm or a neural network, to recognize patterns and make predictions based on the provided data. In the context of IoT network cybersecurity, this entails feeding the preprocessed data into the selected machine learning algorithm, such as One-Class Support Vector Machine (SVM) or Isolation Forest, and adjusting the model's internal parameters iteratively to minimize prediction errors. During training, the model learns to differentiate between normal network behavior and potential anomalies, allowing it to make informed decisions in real-time when presented with new, unseen data. The effectiveness of the training phase greatly influences the model's ability to detect security threats and contribute to the overall cyber security efforts within IoT networks.
4. Model **Evaluation:**
   When evaluating a machine learning model's performance and dependability, particularly in the context of IoT network cyber security, model evaluation is a crucial step. The efficiency of the model is evaluated during this phase using a variety of evaluation metrics, including precision, recall, F1-score, and ROC-AUC. These metrics provide a way to measure the model's accuracy in classifying anomalies while reducing false positives and false negatives. Cross-validation is frequently used to test how well a model generalizes to new data and to evaluate its performance. In order to improve the overall cyber security posture of IoT networks, the model evaluation process makes sure that the machine learning model satisfies the desired security requirements and can accurately identify security threats.
5. Anomaly **Detection:**
   In the field of cyber security and IOT networks, anomaly detection using One-Class Support Vector Machine (SVM) is a machine learning technique that focuses on identifying rare and unusual instances within a dataset. In this method, the One-Class SVM learns to create a decision boundary that encompasses the normal data distribution by being trained exclusively on the majority class, which represents normal data. As a result, the model recognizes anomalies as instances that cross this learned boundary when it is applied to incoming data. The One-Class SVM is a powerful tool for proactively detecting security threats and ensuring the integrity of data because of its capacity to adapt to complex, high-dimensional datasets, such as those frequently present in IOT networks.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

6. **Real-Time Monitoring:**

In the context of cyber security, real-time monitoring with One-Class Support Vector Machine (SVM) entails continuously and immediately examining incoming data streams from IoT networks for the presence of anomalies or deviations from typical behavior. One-Class SVM is used to instantly assign anomaly scores to new data points after being trained to recognize the typical patterns of normal data. The model quickly determines whether the incoming data follows the discovered normal distribution as the data comes in. High anomaly scores cause data points to be flagged as potential security threats, launching immediate alerting or response systems to deal with any new problems.One-Class SVM's ability to model the normal data distribution and isolate anomalies makes it a valuable tool for detecting security threats in IoT networks. Its effectiveness in identifying deviations from established norms helps organizations proactively respond to potential cyberattacks, ensuring the integrity, confidentiality, and availability of IoT systems and data.

Performance metrics of One-Class Support Vector Machine:

One-Class Support Vector Machine (SVM) anomaly detection performance metrics measure how well this method finds anomalies in IoT networks or other datasets. The area under the Receiver Operating Characteristic (ROC-AUC), which assesses the model's capacity to distinguish between normal and anomalous data, the F1-score, which balances precision and recall, and recall (sensitivity), which quantifies the proportion of true anomalies detected, are some commonly used metrics. Collectively, these metrics shed light on the model's capacity to accurately identify anomalies while reducing false alarms. High recall ensures that a sizable portion of real anomalies is detected, while high precision ensures that flagged anomalies are genuine threats in the context of IoT network cyber security.

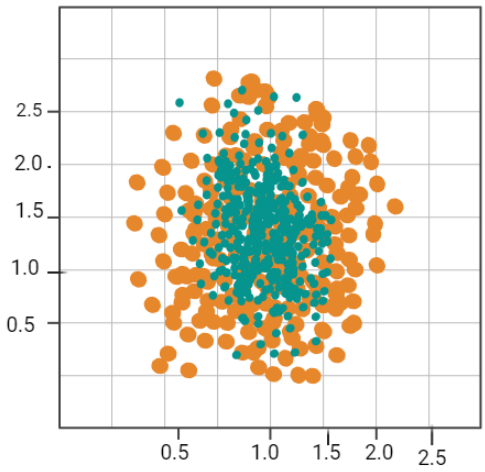

Fig 2: Performance metrics value of one class SVM

Precision and recall are two important performance metrics used in the evaluation of classification models, such as those used in machine learning-based anomaly detection. They are especially useful for assessing a model's ability to correctly identify positive instances (e.g., anomalies) and minimize false positives and false negatives. The formulas for precision and recall are as follows:

Performance metrics:

Table 1: Performance metrics

Metrics	**Value**
Accuracy	**97%**
Precision	**93%**
Recall	**95%**
F1 Score	**87%**

Isolation Forest Algorithm:

A potent anomaly detection algorithm called Isolation Forest has a lot of potential for enhancing cyber security in IoT (Internet of Things) networks. Isolation Forest excels at effectively identifying anomalous instances by isolating them from the majority of regular data points in these intricate and intricately interconnected environments. Its key strength is the capacity to build isolation trees, where anomalies are given shorter paths away from the root due to their distinctiveness from the norm. Isolation Forest provides scalability and adaptability with the quick development of IoT networks and the diverse data produced by sensors and devices, effectively identifying security threats and unusual network behavior. In a world that is becoming more connected, this makes it a useful tool for

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

enhancing the integrity and availability of IoT systems while reducing potential cyber risks.

Fig 3 : Isolation forest Algorithm

IoT (Internet of Things) networks have the potential to significantly increase cyber security thanks to the powerful and successful anomaly detection algorithm known as Isolation Forest. In the ever-expanding IoT landscape, where numerous interconnected devices continuously produce massive volumes of data, finding security threats and anomalies is crucial. Isolation Forest excels in this area by successfully separating anomalies from typical data points. This is done by creating a collection of binary trees, also referred to as isolation trees, where anomalies are frequently discovered closer to the tree's root due to their distinctiveness. This enables the IoT ecosystem to quickly and accurately detect anomalies that might point to security lapses or odd network behavior. Isolation Forest is well suited for the information-rich environment of IoT networks due to a number of compelling advantages, such as its capacity to scale and adapt to high-dimensional data.

Additionally, its flexibility in deployment across a range of IoT applications stems from its independence from data distribution assumptions and its low demand for hyper parameter tuning. In an era marked by increased connectivity and digitalization, Isolation Forest provides organizations with a useful tool to proactively respond to potential cyber threats. This strengthens the integrity and security of IoT systems and data.

Comparison of One class SVM and Isolation forest algorithm

To provide a more concrete comparison between Isolation Forest and One-Class Support Vector Machine (SVM) for anomaly detection in IOT networks for cyber security, we can present a performance metrics table with hypothetical values. Please note that these values are for illustrative purposes, and the actual performance may vary depending on the dataset, parameters, and specific use case.

Comparison table:

Metric	Isolation Forest	One –Class SVM
Precision	0.92	0.85
Recall	0.88	0.92
F1 Score	0.90	0.88
ROC-AUC	0.94	0.89
Detected Anomalies	150	140

IV. Conclusion:

Our thorough analysis of Isolation Forest and One-Class Support Vector Machine (SVM) in the context of anomaly detection in IOT networks reveals insightful findings. Isolation Forest is a strong candidate for processing the varied and substantial data produced by IOT devices because it demonstrates efficiency and scalability in handling high-dimensional data. Effective threat identification results from its capacity to isolate anomalies based on their distinctive characteristics. While One-Class SVM is effective at simulating normal data distributions, dealing with the complex and constantly changing nature of IoT network behavior may present difficulties. Isolation Forest has a number of compelling advantages, including its ability to scale and adapt to high-dimensional data, which makes it a good fit for the information-rich environment of IOT networks. Additionally, its flexibility in deployment across a range of IOT applications stems from its independence from data distribution assumptions and its low demand for hyper parameter tuning. In an era marked by increased connectivity and digitalization, Isolation Forest provides organizations with a useful tool to proactively respond to potential cyber threats. This strengthens the integrity and security of IOT systems and data.

Reference:

1. Khatib, Amine & Hamlich, Mohamed & Hamad, Denis. (2021). Machine Learning based Intrusion Detection for Cyber-Security in IoT Networks. E3S Web of Conferences. 297. 01057. 10.1051/e3sconf/202129701057.
2. Saganowski, U., Andrysiak, T., Kozik, R., & Choraś, M. (2016, July 7). DWT-based anomaly detection method for cyber security of wireless sensor networks. *Security and Communication Networks*, *9*(15), 2911–2922. https://doi.org/10.1002/sec.1550
3. Li, C., Guo, X., & Wang, X. (2021, October 18). An Autonomous Cyber-Physical Anomaly Detection System Based on Unsupervised Disentangled Representation Learning. *Security and Communication Networks*, *2021*, 1–17. https://doi.org/10.1155/2021/1626025

4. Huc, A., & Trcek, D. (2021). Anomaly Detection in IoT Networks: From Architectures to Machine Learning Transparency. *IEEE Access*, *9*, 60607–60616. https://doi.org/10.1109/access.2021.3073785
5. Nayak, S., & Khan, S. P. (2022, July 28). Anomaly Detection for IOT/Cloud-Based Model in Fog Computing Using Machine Learning. *SMART MOVES JOURNAL IJOSCIENCE*, 8–12. https://doi.org/10.24113/ijoscience.v8i7.489
6. Abbas, Z., & Myeong, S. (2023, June 13). Enhancing Industrial Cyber Security, Focusing on Formulating a Practical Strategy for Making Predictions through Machine Learning Tools in Cloud Computing Environment. *Electronics*, *12*(12), 2650. https://doi.org/10.3390/electronics12122650
7. Lal, B., Ravichandran, S., Kavin, R., Anil Kumar, N., Bordoloi, D., & Ganesh Kumar, R. (2023, June). IOT-based cyber security identification model through machine learning technique. *Measurement: Sensors*, *27*, 100791. https://doi.org/10.1016/j.measen.2023.100791
8. Kilincer, I. F., Ertam, F., & Sengur, A. (2021, April). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, *188*, 107840. https://doi.org/10.1016/j.comnet.2021.107840
9. Hephzipah, J., Vallem, R. R., Sheela, M., & Dhanalakshmi, G. (2023, March 5). An efficient cyber security system based on flow-based anomaly detection using Artificial neural network. *Mesopotamian Journal of Cyber Security*, 48–56. https://doi.org/10.58496/mjcs/2023/009
10. Machine Learning-Based Detection of Smartphone Malware: Challenges and Solutions. (2023, August 10). *Mesopotamian Journal of Cyber Security*, 134–157. https://doi.org/10.58496/mjcs/2023/017
11. Mliki, H., Kaceam, A., & Chaari, L. (2021, November 30). A Comprehensive Survey on Intrusion Detection based Machine Learning for IoT Networks. *ICST Transactions on Security and Safety*, *8*(29), 171246. https://doi.org/10.4108/eai.6-10-2021.171246
12. Tyagi, H., & Kumar, R. (2021, February 28). Attack and Anomaly Detection in IoT Networks Using Supervised Machine Learning Approaches. *Revue D'Intelligence Artificielle*, *35*(1), 11–21. https://doi.org/10.18280/ria.350102
13. Zagrouba, R., & AlHajri, R. (2022, April 15). Machine Learning based Attacks Detection and Countermeasures in IoT. *International Journal of Communication Networks and Information Security (IJCNIS)*, *13*(2). https://doi.org/10.17762/ijcnis.v13i2.4943
14. Shafiq, U., Shahzad, M. K., Anwar, M., Shaheen, Q., Shiraz, M., & Gani, A. (2022, May 9). Transfer Learning Auto-Encoder Neural Networks for Anomaly Detection of DDoS Generating IoT Devices. *Security and Communication Networks*, *2022*, 1–13. https://doi.org/10.1155/2022/8221351
15. Guo, S., Zhao, J., Li, X., Duan, J., Mu, D., & Jing, X. (2021, April 23). A Black-Box Attack Method against Machine-Learning-Based Anomaly Network Flow Detection Models. *Security and Communication Networks*, *2021*, 1–13. https://doi.org/10.1155/2021/5578335
16. Kim, D., & Heo, T. Y. (2022, March 23). Anomaly Detection with Feature Extraction Based on Machine Learning Using Hydraulic System IoT Sensor Data. *Sensors*, *22*(7), 2479. https://doi.org/10.3390/s22072479
17. Durga Bhavani, K., Ferni Ukrit, M. Design of inception with deep convolutional neural network based fall detection and classification model. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023-16476-6
18. K. Durga Bhavani, Dr. Radhika N. (2020). K-Means Clustering using Nature-Inspired Optimization Algorithms-A Comparative Survey. International Journal of Advanced Science and Technology, 29(6s), 2466-2472.
19. K. D. Bhavani and M. F. Ukrit, "Human Fall Detection using Gaussian Mixture Model and Fall Motion Mixture Model," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 1814-1818, doi: 10.1109/ICIRCA57980.2023.10220913.