Sybil Attack Detection in Industrial Internet of Thing Network using Hybrid Model

Adnan Nadeem
Faculty of Computer and Information System
Islamic University of Madinah
Madinah 42351, Saudi Arabia

Amir Mehmood
Department of Computer Science
Faculty of Computing
Al-Kawthar University
Karachi 75300, Pakistan

Mohammad Zubair Khan
Faculty of Computer and Information System
Islamic University of Madinah
Madinah 42351, Saudi Arabia

Muhammad Ashraf
Department of Physics,
Federal Urdu University of Arts, Science & Technology
Karachi 75300, Pakistan

Syed Saood Zia
Department of Computer Science
Faculty of Computing
Al-Kawthar University
Karachi 75300, Pakistan

Shakir Karim Baksh
Department of Computer Science
Faculty of Computing
Al-Kawthar University
Karachi 75300, Pakistan

Hani Almooamari
Faculty of Computer and Information System
Islamic University of Madinah

***Abstract*— The Industrial Internet of Things (IIoT) describes to the incorporation of the Internet of Things (IoT) with artificial intelligence (AI) in automating processes in sophisticated industrial things and supply chains. IIoT facilitates smart manufacturing, asset tracking, supply chain optimization, energy management, and remote monitoring and control. However, these IIoT applications are vulnerable to various cyberattacks. This paper focuses on the sybil attack, in which an attacker within a distributed system uses multiple fake identities to achieve their objectives. To study this, we meticulously generated a dataset of 100 node attribute instances by simulating an IIoT environment. We then performed preprocessing, including applying the SMOTE technique to address class imbalance. Initially, we implemented four baseline models and compared their performance with our proposed hybrid model combining GNN and transformer models. The simulation results demonstrate high accuracy of proposed model in detecting the sybil attacks.**

Keywords*— *Industrial Internet of Things (IIoT), Sybil Attack, Performance Analysis, Sybil attack detection

I. Introduction

These years industrial revolution witness the possibilities of the application of Internet of Thing (IoT) with artificial intelligence in industrial process of manufacturing and supply chain. The incorporation of IoT and AI which is often termed as Industrial IoT (IIoT) in industrial scenarios enable them to automate, improve and automatically manage the industrial processes. The main concept behind the fourth and fifth business revolution i.e. industry 4.0, industry 5.0 was the digitalization of industrial process. IIoT emerge as one of the suitable technology to achieve industry 4.0/5.0 objectives in terms of smart manufacturing, real-time monitoring & control, improve decision-making and optimizing operations. As an example, smart manufacturing using IIoT has been proposed in [1] to automate and optimize the process. Several other applications [3] of IIoT has already been proposed in recent years.

However, along with these advantages of IIoT applications in industrial scenarios it also poses vulnerabilities against various cyber-attacks [2]. The cyber attacks ranges from denial of services, unauthorize access, sybil attack, integrity violation and data stealing. In industrial IoT its critical to protect against most common cyber threat to ensure smooth operations and processing of daily task. Considering the importance of protecting IIoT against cyber threats , in this we study the sybil attack. This attack may keep false identities to change decision and control networks.

The sybil attack has been investigate and solution are proposed in IoT. For example, authors in [24] propose detecting sybil attack in IoT using edge computing, in [25] authors have employed trust mechanism to deal with sybil attack and IoT and, transfer learning combined with Game theory approach was proposed in [4] to detect sybil attack in IoT. However, we observe no significant work has been done to deal with sybil attack in IIoT. So, in this paper, we propose and evaluate performance of sybil attack detection mechanism in IIoT scenario. We first employed four baseline models on the dataset [available at github with the title of Industrial-Internet-of-Things-IIoT-Dataset-for-Sybiil-Attack] to compare their performance and then applied propose Hybrid GNN (Graph Neural Network) plus Transformer model which results in a very high accuracy.

Section 2 briefly describe the related works while section 3 use proposed methodology including details of each step. Experimental results and discussion are presented in Section 4. In final section conclude our study and also focus on future research can be done.

II. Related Work

The sybil attack is one of the most advanced threats that can happen to wireless sensor networks. In the theater of network subversion, rogue nodes fabricate counterfeit personas to erode the very sinews of system integrity. This study examines various methodologies for employing machine learning to identify malicious assaults. We meticulously curated the dataset by simulating IIoT scenarios using Python that has 100 network instances, each characterized by essential attributes including

Identify applicable funding agency here. If none, delete this text box.

fluctuating RSSI values, variable transmission frequencies, inconsistent message intervals, and packet losses. Here, the pursuit is not merely detection, but discernment: the unveiling of deception woven into the digital mesh, where every anomaly whispers the trace of a concealed adversary.

The study is important because it could help with network security problems in the real world. There are more IoT devices and wireless networks than ever before, so we need detection systems that work well and are reliable. A sybil attack is a kind of digital identity theft in which one malicious user can act like many real users on a network.

The IIoT leverages intelligent devices that communicate with one another via the internet [3]. IoT environments have a lot of smart gadgets that can send, receive, gather, and process data from one another [5]. These smart devices that are connected to each other enable us monitoring environmental conditions and exercising precise control over settings are key functions [6]. The annual economic impact of IoT technology is projected to reach USD 11.1 trillion by 2025 [7]. The widespread adoption of consumer-focused IoT systems has, in turn, fostered the integration of this technology into diverse industrial applications, thereby giving rise to IIoT technology [8]. The IIoT constitutes a framework that utilizes interconnected intelligent devices within an industrial setting to connect various components—including actuators, sensors, controllers, and sophisticated control systems—for the purposes of data analysis and the optimization of industrial processes. thereby improving execution speed, reducing costs, and facilitating dynamic management of the industrial environment [9].The rapid expansion of IIoT networks has resulted in a growing array of possible security weaknesses, rendering intrusion detection a critical field of study. The available Intrusion detection methods like signature based, anomaly based faced big challenges for securing IIoT system because of their active features and constrained resources [10]. Therefore, machine learning (ML), neural networks (NN), and genetic algorithms (GA) have become increasingly important in improving intrusion detection system (IDS), refining feature selection, and optimizing system designs.

[11-12]. Because of the characteristics of features of Internet of things, like minimum reassures, distributed system, a special purpose intrusion detection method is needed. Deep learning methods have been used to detect anomaly behavior by classifying network traffic. The complex nature and high dimensionality of IoT data need the use of feature selection methods to reduce computational complexity while maintaining detection accuracy [13]. Various machine learning methods e.g. Decision Tree, Random Forest, K-nearest neighbors and SVM were used for IDS though the dataset was labeled [14]. Despite these results, the challenges occur due to high dimension dataset, to resolve this challenge utilize the new feature selection algorithms to optimize input data while keeping essential features. Advanced machine learning models like ensemble learning and deep learning methods should be used to resolve this challenge. [15]. The genetic algorithm method can be use for tackle the issue of feature selection in IDS system. Genetic algorithms incorporate crossover alteration and selection of best features therefore reducing dimensionality of dataset improving accuracy research has been proved the accuracy of Genetic algorithms for feature selection in IDS systems.

In [16] authors propose notable enhancements in detection accuracy by the application of Genetic Algorithms to optimize characteristics and the utilization of Random Forest for classification. In this article [17] incorporates genetic algorithms method to select best features sets in wireless sensor networks, use in the Internet of Things, shows more accurate intrusion detection as compare to available legacy IDS systems. Authors used here genetic algorithm with SVM for feature optimization, results show increase in classification accuracy by shrinking repeated features.

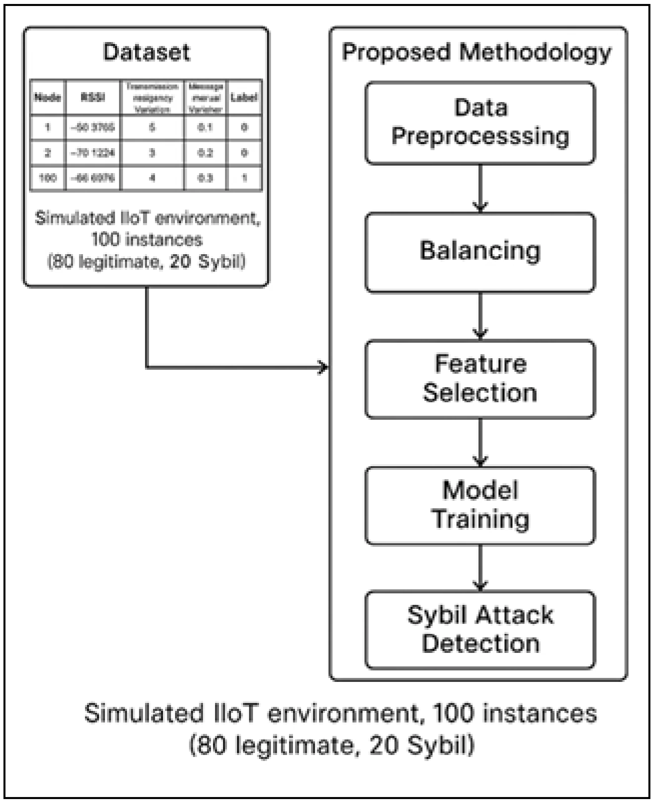
In legacy, IDS methods in reference to Industrial Internet of things heavily depends on machine learning algorithms like Decision Tree, Support vector method and KNN [18, 19].

These algorithms are easy to use in terms of running cost and understanding but they are near about fail when dataset is multi-dimensional, and also Machine learning techniques need a lot of feature selection engineering some times do not works when threat become complex [20]. Now days progress in deep learning methods like as Conventional Neural Networks and Long short Term memory (LSTM) networks shows some better results in intrusion detection system for IIoT. [21, 22]. But these models requires more resources like memory and time, which means they need a lot of processing power. Due to this it is hard to use them on IIoT devices that don't have a lot of resources which leads it as costly system[23]. Also, deep learning models are likely to overfit if they don't have the right regularization and hyperparameter tuning.

III. Proposed Methodology

Our proposed sybil attacks detection approach in IIoT systems is shown in Figure 1. For simplicity our approach can be employed in a series of steps. The most important is the dataset which has relevant features for the detection of the attack. Data preprocessing is the next step which involves cleaning the data, and then run it through a series of steps such as balancing the data, choosing the most meaningful features, and training a model. These steps matter because real-world data often has problems, some classes show up far more than others, models may easily overfit, and not every recorded feature is actually useful.

The dataset contains 100 samples—80 from legitimate nodes and 20 from sybil attacker nodes. The dataset consists of useful details about each device. Each device has a Node ID, serving as its distinct identifier, the RSSI (Received Signal Strength Indicator), which shows signal power (e.g., 60.27 dBm), the Transmission Frequency, which means how often each device sends a message, the Message Interval Variation, which captures irregularities in how frequently nodes communicate, the Packet Drop Rate, a measure of how many messages fail to get through, and also includes a label that indicates whether the node is legitimate (0) or a sybil attacker node (1). The overall method is broken down into a series of steps, which are illustrated in the Fig. 1.



Node	RSSI	Transmission requency Variation	Message interval Variation	Label
1	−50.3765	5	0.1	0
2	−70.1224	3	0.2	0
100	−66.6076	4	0.3	1

Fig. 1. Proposed approach for Detecting the sybil attacks in IIoT Networks

A. Data Preprocessing

We started off with the dataset in a CSV file (sybil_detection_dataset.csv). A few entries had missing values, so instead of leaving gaps, we just filled them with the mode since it made the data more consistent. Some records were repeated, and a couple of RSSI values didn't make sense, so those were fixed or removed.

After that, histograms are plotted for the main numeric features (RSSI, Transmission Frequency, Message Interval Variation, Packet Drop Rate) for gaining a preliminary understanding about how the values were spread out. Then box plots generated were to identify any outliers that might be present.

Finally, a correlation matrix was constructed to check if things like RSSI had any connection with Packet Drop Rate.

Since Node ID was a text-based category, it is converted in to numbers using ordinal encoding. That way, the machine learning models could actually work with it, but it could still be traceable back which device was which.

The problem with the features was that they were all measured on completely different scales. For example, RSSI values are negative, while packet drop rates are percentages. To make them comparable, a min–max scaling was applied, which adjusts all values to fit within the range of 0 to 1. This prevents any single feature from dominating the others just because of its numerical size.

Once that was done, the dataset was split into two parts. The first part (X) included the features RSSI, transmission frequency, message interval variation, packet drop rate, and Node ID, while the second part (Y) was the label indicating whether a node was normal or sybil attacker. To make sure that both categories were represented fairly, a stratified train–test split was applied. This left roughly 80% of the records for training and 20% for testing, ensuring a balanced mix of normal and sybil nodes in both sets.

B. Handling Class Imbalance with SMOTE

While working with the dataset, we found that the number of legitimate nodes was much higher than the number of sybil nodes. If left unaddressed, this imbalance could make the model in favor of higher number of class and overlook the sybil attacks. To reduce this risk, the Synthetic Minority Over-sampling Technique (SMOTE) applied on the training data. Instead of simply repeating the minority samples, SMOTE produces new ones by filling in values between existing points, which adds some variety to the dataset. Once this was applied, the training set ended up balanced with 64 samples for each class. The test set was left unchanged, keeping its 16 samples of each type so that the evaluation stayed realistic. To check the effect, we looked at the class distributions before and after SMOTE.

C. Selecting Features

With the dataset balanced, the next task was to decide which features were most useful for classification. The attributes under consideration included RSSI, transmission frequency, message interval variation, packet drop rate, and Node ID. We relied on two complementary strategies.

1) Pearson Correlation Coefficient: We first examined the association between features and the target label using Pearson correlation. To better visualize the patterns, the results were presented in a heatmap. For example, packet drop rate and message interval variation showed a noticeable association with the class label. Using a threshold of 0.15, we kept the features that demonstrated stronger correlations.

2) Genetic algorithm search: Since correlation only uncovers linear patterns, we also turned to a genetic algorithm (GA). The method involved testing different combinations of features and progressively retaining those that improved classification performance. This process allowed the genetic algorithm to reveal relationships between variables that correlation analysis by itself could not identify. The features it selected were then used in the model training phase.

D. Training and Testing the Model:

We trained three traditional classifiers such as Logistic Regression, AdaBoost, and Random Forest on the balanced, chosen features. We used grid search to fine-tune the hyperparameters, such as C for Logistic Regression and n number of estimators for ensembles. A GNN was utilized to leverage the IIoT network topology. The proposed hybrid model GNN plus Transformer is utilized to find both structural and sequential features in industrial internet of things networks traffic, TensorFlow and Keras are utilized to figure out the Deep Learning model, which has many dense layers and dropout for regularization. This wide range of models makes sure that performance can be compared fully across different approaches. In this case, nodes stand for devices (through Node ID), and edges are based on RSSI and Transmission Frequency. The

GNN combined features to classify nodes. The proposed hybrid model is examined based on accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC-ROC). For each Algorithm, we created confusion matrices and classification reports, and we used ROC curves to evaluate how effectively the models distinguished between normal and the sybil attacker nodes. By combining careful data preparation, balanced classes, and advanced modeling, this approach provides a reliable way to detect the sybil nodes while capturing the relationships specific to IIoT networks.

IV. Experimental Results

A. Dataset Overview and Preprocessing Methodology

The dataset of 100 network nodes, each described by five features: an identification number, signal strength, data transmission frequency, variation in message timing, and the rate of dropped messages. In the initial analysis, we found an imbalance in the data—80 nodes were labelled as normal, while only 20 were labelled as malicious as shown in Table 1.

TABLE I. Initial Analysis of 100 Nodes with Five Features

Node_ID	RSSI	Transmission_Frequency	Message_Interval_Variation	Packet_Drop_Rate	Label
Node_1	-59.006572	7.169259	0.535779	0.005855	0
Node_2	-62.276529	9.158709	0.550878	0.007139	0
Node_3	-58.704623	9.314571	0.60835	0.013763	0
Node_4	-56.59349	8.935445	0.60538	0.013052	0
Node_5	-60.448307	9.674429	0.362233	0.009895	0
…	…	…	…	…	…
…	…	…	…	…	…
Node_96	-60.371757	26.963587	1.361418	0.176541	1
Node_97	-59.58149	20.580713	1.67902	0.114343	1
Node_98	-59.484672	25.766029	1.56146	0.17647	1
Node_99	-59.3971	25.290614	1.662527	0.214093	1
Node_100	-60.117294	19.285149	1.629526	0.261891	1

This asymmetry mimics reality as most of the time, the sybil attacks, despite their destructive capability, comprise a tiny fraction of the network stream. It creates problems for machine learning systems, which generally assist the majority class.

B. Feature Engineering and Scaling Techniques

With the application of feature engineering and scaling techniques, any machine learning pipeline would benefit from balancing according to the scaling daisy chain. Actually, the processes would be applied in the order of processing, beginning with the Min-Max Scaler, as shown below, supplied with an appropriate ordinal encoding class, etc.

Every machine learning classifier is equipped with techniques of quantitative feature scaling. Whenever a quantitative feature is represented numerically, the feature is subjected to a quantitative transforming function called a scaling function, which bounds the feature to the [0,1] interval as shown in Table 2. Comparing currencies of different countries becomes impractical and irrational. However, if we normalize and standardize the currencies, a $ 100 bill and a 10,000 yen note become rational to compare.

TABLE II. Applied Scaling Techniques for 100 Network Nodes Across Five Features

Node_ID	RSSI	Transmission_Frequency	Message_Interval_Variation	Packet_Drop_Rate	Label
0	0.696879	0.043548	0.220574	0.016035	0
1	0.55489	0.129598	0.217169	0.015254	0
2	0.370639	0.13634	0.2477	0.033715	0
3	0.926376	0.096594	0.24533	0.035063	0
4	0.534348	0.152035	0.106447	0.025118	0
…	…	…	…	…	…
…	…	…	…	…	…
95	0.503993	0.89812	0.67717	0.597087	1
96	0.602362	0.623639	0.819535	0.373233	1
97	0.604041	0.848035	0.914032	0.719022	1
98	0.586093	0.827378	0.849186	0.658011	1
99	0.572693	0.567602	0.828254	0.890336	1

C. Class Imbalance Handling with SMOTE Implementation.

Understanding the Original Data Distribution in class imbalance problems, SMOTE is an approach to synthesize new minority class observations in the training data. Descriptive statistics indicate the original dataset of super nodes and attacker was polarized with a 4 to 1. An overlying dominant class can mean that the subordinate class is overlooked or ignored. This is called the accuracy paradox in situations when we discern superordinate dominant class labels.

D. Application and Validation of SMOTE.

To tackle this imbalance problem, the SMOTE technique was applied, and synthetic examples of the minority class until each class contained 80 instances was created. With SMOTE, all the classes begin and end with 80 instances, as the post-processing visualization clearly shows in Fig. 2.

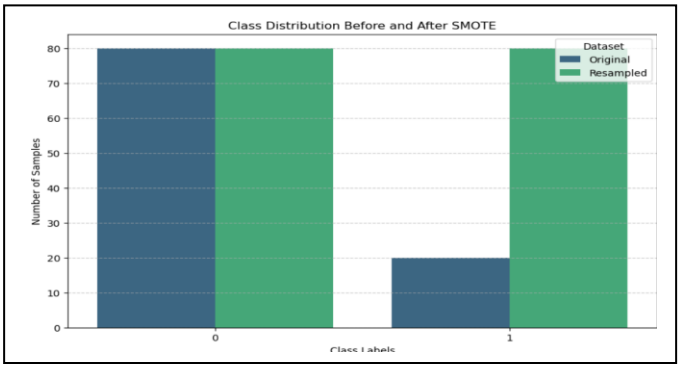


Fig. 2. Class Data Distribution before and after SMOTE

SMOTE creates synthetic examples in the ‘spaces’ along the line segments joining the instances of the minority class and their k-nearest neighbors. SMOTE, along with the minority class, is able to fill in the ‘spaces’ within the synthetically created neighborhoods, as SMOTE is able to increase the class density as shown in Fig. 3.

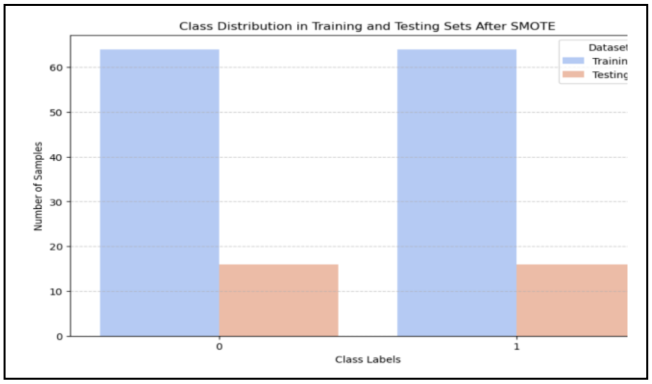


Fig. 3. Class Data Distribution in training and testing sets after SMOTE

I. *Feature Selection:* To find the most significant features that contribute to classification, feature selection is carried out. The suggested model ranks features according to their relevance using Decision Tree feature importance. In order to reduce dimensionality and increase computational efficiency, the top 4 out 10 features are chosen for training. By doing this, the model 541's predictive accuracy is increased by ensuring that it concentrates on the most informative features. The model improves generalization and training times by removing 542 superfluous features. Figure 4 displays the feature selection for the sybil attack dataset. For features selection research used the Decision Tree and Pearson correlation coefficient algorithms.

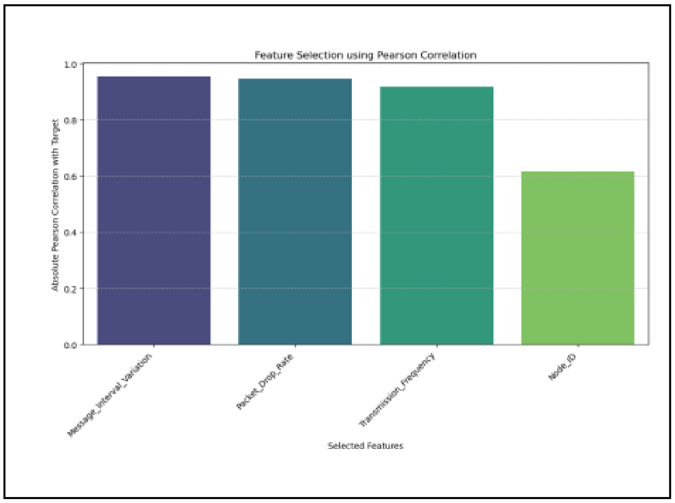


Fig. 4. Feature selection

Both structural and sequential relationships in IIoT traffic are intended to be captured by the hybrid GNN + Transformer model. In this hybrid model it uses two hidden layers and sixty-four hidden nodes with a transformer encoder of two layers, as well as four attention heads and also a forty-eight projection dimension. For binary categorization, the result from hybrid components is concatenated and run by a linear layer with sigmoid activation function, uses binary cross entropy, same time graph edges are constructed using KNN method.

Lastly, we use accuracy, precision, recall, and F1-score to assess the model on test data that hasn't been seen yet. We also make sure that the training and test accuracy curves are watched for overfitting.

II. *Experiment:* Binary classification—the process of distinguishing between attacks and benign traffic—was the main focus of the experiment. AdaBoost, Random Forest (RF), Logistic Regression (LR), and our suggested GNN–Transformer hybrid model were the four baseline models that were previously introduced. We were able to determine how well traditional and hybrid approaches handle simple separation tasks by using this first test as a benchmark.

III. *Performance Metrices:* We assess each ML model using the following metrics.

a. Accuracy: The accuracy metric calculates the percentage of accurate predictions made by all samples in a dataset in order to assess the classification model.
Acc = (TP+TN)/(TP+TN+FP+FN)

b. Recall: The ratio of recognized classes to the total number of instances of a specific class is known as recall.
Rec = (TP)/(TP+FN)

c. Precision: The ratio of correctly classified classes to all positive classifications is known as precision.
Pre = (TP)/(TP+FP)

d. F1-Score: The F1 calculates the mean of recall and precision in the following way:
F1 = (2) * (Pre * Rec)/(Pre + rec)

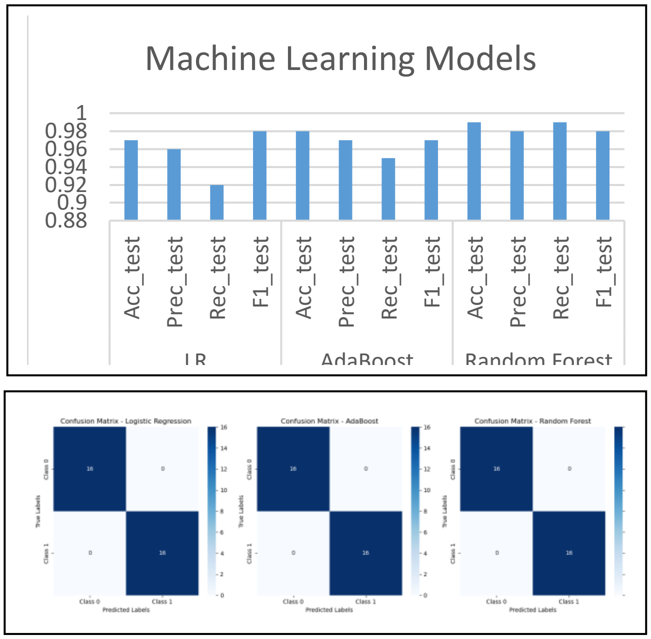


Fig. 5a. Comparison of baseline models LR, RF and Adaboost (b) Confusion Matrices of baseline models LR, RF and Adaboost

IV. *Discussion On Results:* In Figure 5, we compare the results of the experiment between the selected baseline models, LR, RF, and AdaBoost. High detection accuracy for the anomalies in the sybil attack dataset was attained by all three models. Every model performs exceptionally well, with 99% for RF, AdaBoost 98% and LR 97% accuracy. As our propose Hybrid GNN+ Transformer model perform well with 100% accuracy shown in Fig. 5a, confusion matrices in Fig. 5b.

Our hybrid suggested GCN + Transformer model obtained 100% accuracy, with a precision of 100%, recall of 100%, and F1 score of 100% as shown in Figure 6.

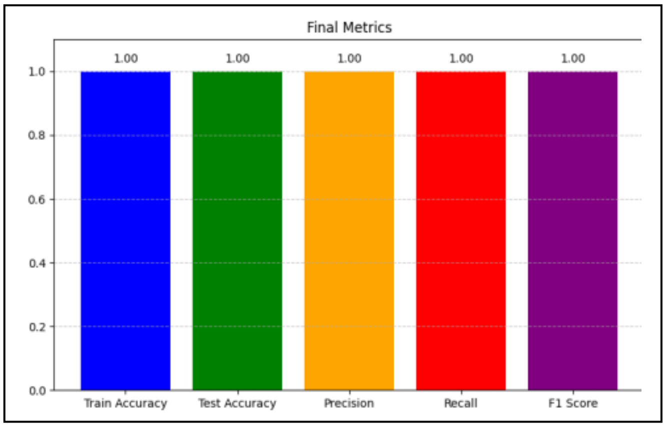


Fig. 6. GNN + Transformer Model Train and Testing Accuracies

V. Conclusion

Given the importance of security in advanced industrial applications that utilize technology, this research proposes a method for detecting the sybil attacks in Industrial Internet of Things (IIoT) networks. We began by implementing baseline models using a dataset with five features to assess their performance in detecting the sybil attacks. Following this, we introduced a hybrid model that combines Graph Neural Networks (GNN) with a transformer model. The experimental results demonstrate high accuracy in the tested scenarios. However, the study is limited to the dataset used in our experiments. In the future, we plan to apply our proposed approach to other benchmark datasets to evaluate its scalability and adaptability, along with the main performance metrics.

References

[1] L. L. Yadla, R. A. Mudragada, H. Poka and T. Vignesh, "Smart Manufacturing In Industries Using Internet Of Things," *International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2023, pp. 1-5.

[2] Max Pilates, "Industrial IoT security: top concerns & actionable strategies", published by itransition on their webportal , Dec 2023.

[3] Chalapathi, G.S.S. Chamola, V.; Vaish, A. Buyya, R, "Industrial internet of things (iiot) applications of edge and fog computing: A review and future directions". Published in Fog/Edge Computing For Security, Privacy, Applications Springer Book, Switzerland, 2021; pp. 293–325.

[4] Ullah F, Turab A, Ullah S, Cacciagrano D, Zhao Y, "Enhanced Network Intrusion Detection System for Internet of Things Security Using Multimodal Big Data Representation with Transfer Learning and Game Theory". Sensors (Basel). 2024 Jun 26;24(13):4152.

[5] Shishehgarkhaneh, M.B.; Moehler, R.C.; Moradinia, S.F. "Blockchain in the Construction Industry between 2016 and 2022: A Review Bibliometric, and Network Analysis". Publised in MDRP Smart Cities 2023, 6, 819–845.

[6] Ahmad, T. and Zhang, D, "Using the internet of things in smart energy systems and networks". Publised in Journal of Sustainable Cities and Society. 2021, Vol.68, 102783.

[7] A. Nistor, and E. Zadobrischi, "Analysis and Estimation of Economic Influence of IoT and Telecommunication in Regional Media Based on Evolution and Electronic Markets in Romania". MDPI *Telecom* **2022**, *3*, 195-217.

[8] E. Sisinni, A. Saifullah, S. Han, U. Jennehag and M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," in *IEEE Transactions on Industrial Informatics*, vol. 14, pp.4724-4734, 2018.

[9] H. Xu, W. Yu, D. Griffith and N. Golmie, "A Survey on Industrial Internet of Things: A Cyber-Physical Systems Perspective," in *IEEE Access*, vol. 6, pp. 78238-78259, 2018.

[10] Yao W, Shi H, Zhao H. "Scalable anomaly-based intrusion detection for secure internet of things using generative adversarial networks infog environment". Journal of Network and Computer Applications. Vol .214: May 2023, 103622.

[11] J.L Silva, R.Fernandes, N.Lopes, "Performance Study on the Use of Genetic Algorithm for Reducing Feature Dimensionality in an Embedded Intrusion Detection System". MDPI *Systems* **2024**, *12*, 243.

[12] Xingjuan FA, Hui LI, Xinglong LI, Fangtong GU. Illegal intrusion detection of internet of things based on deep mining algorithm. TechnicalGazette. 2023.

[13] Kalimuthu VRVK. "Modeling of intrusion detection system using double adaptive weighting arithmetic optimization algorithm with deeplearning on internet of things environment". Published by Brazilian Archive of Biology and Technology. Vol.67, 2024.

[14] Al-Saleh A. "A balanced communication-avoiding support vector machine decision tree method for smart intrusion detection systems". Published by Scientific Report. 2023;13:9083.

[15] Yang J, Wang H, Guo S, Wang L. "Deep learning based intrusion detection for iot networks". IEEE Trans Ind Inf. 2020;16(11):7173–81.

[16] Mukherjee S, Sharma M, Sahoo B. "Genetic algorithm based feature selection and detection of intrusion in IoT". Proc Comput Sci.2018;132:284–92.

[17] Pires P, Govindarajan A, Queiroz C. "Enhancing security in wireless sensor networks using genetic algorithms for anomaly detection". Sensors. 2021;21(5):1234.

[18] Buczak AL, Guven E. "A survey of data mining and machine learning methods for cyber security intrusion detection". IEEE Communication Survey and Tutorial. 2016.

[19] Aljawarneh MBYM. Shadi & Aldwairi: "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model". Journal of Computer Science. 2017.

[20] A. Khacha, Y.H. R. Saadouni, Aliouat, Z. "Hybrid deep learning-based intrusion detection system for industrial internet of things". 5th International Symposium on Informatics and its Applications (ISIA), M'sila, Algeria (2022).

[21] C.S Dash et.al., "An optimized lstm-based deep learning model for anomaly network intrusion detection". Scientific Report (2025).

[22] C. Yin, Y. Zhu, J. Fei and X. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," in *IEEE Access*, vol. 5, pp. 21954-21961, 2017

[23] N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41-50, Feb. 2018.

[24] Yan, J., Jiang, T., Lin, L. *et al.* "A novel Sybil attack detection scheme in mobile IoT based on collaborate edge computing". *J Wireless Com Network* **2023**, 25 (2023).

[25] J. Hassan, A. Sohail, A. Ismail Awad, M. A. Zaka,LETM-IoT: "A lightweight and efficient trust mechanism for Sybil attacks in Internet of Things networks",Ad Hoc Networks,Vol.163,2024.