Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 8
August 2023

Malicious Node Detection in a Wireless Sensor Network

[1]Anmol, [2]Pallavi Joshi

[1]Research Scholar, Department of Computer Science, Amrita Vishwa Vidyapeetham, Mysuru

[2]Assistant Professor, Department of Computer Science, Amrita Vishwa Vidyapeetham, Mysuru

Abstract

Wireless Sensor Networks (WSNs) are vulnerable to malicious attacks that can disrupt operations and compromise data integrity. This paper proposes a method for detecting malicious node in WSNs using the Cooja network simulator and machine learning (ML) algorithms, specifically Random Forest and SVM. The approach involves collecting a dataset of normal and malicious traffic patterns simulated in Cooja. Features are extracted from the network traffic data, including packet size, addresses, timing, frequency, and network topology. Feature selection techniques identify informative features for distinguishing between normal and malicious node. The dataset splits into testing and training sets, and the Random Forest algorithm is trained using the training set. Performance evaluation measures accuracy, precision, recall, and F1-score. Additionally to enhance detection performance further, the SVM algorithm is incorporated. Known for its ability to handle high-dimensional data and separate complex decision boundaries, SVM constructs a hyperplane for effective identification of malicious nodes inthe WSN.. The optimized models are deployed in real-time WSN environments to monitor incoming traffic continuously. Alerts are generated upon detecting malicious node, enabling prompt response and mitigation. This proposed method offers an effective means of detecting malicious node, improving the security and reliability of WSNs. The results highlight the potential of machinelearning algorithms, specifically SVM and Random Forest, inaccurately classifying and identifying malicious patterns. By incorporating these techniques, robust security mechanisms for WSNs can be developed.

Keywords: - Wireless Sensor Networks (WSNs), Malicious node detection, Cooja network simulator, Machine learning algorithms, SVM, Random Forest, Dataset collection, Feature extraction, Feature selection, Training and testing sets, Performanceevaluation, Hyperparameter tuning, Real-time monitoring, Alert generation, Data integrity, Network security, Robust security mechanisms.

I. Introduction: -

WSNs have come to the force as a vital technology in various domains, including environmental monitoring, healthcare, agriculture, and infrastructure management. These networks consist of many small, resource-constrained sensor nodes that communicate wirelessly to collect and transmit data. However,the open and distributed behavior of WSNs causes malicious attacks that can compromise data integrity, disrupt network operations, and compromise the privacy of sensitive information.To address these security challenges, this paper focuses on the detection of malicious nodes in WSNs using the Cooja network simulator and machine learning algorithms. The objective is to develop an efficient and effective method to identify and classify abnormal node behaviour associated with attacks. The proposed approach leverages the Cooja network simulator, which provides a realistic environment for simulating WSNs and generating diverse network traffic scenarios. By simulatingboth normal and malicious traffic, a comprehensive dataset can be collected for training and evaluating machine learning models.

Machine learning algorithms, specifically Support Vector Machines (SVM) and Random Forest, are employed to learn patterns and characteristics that differentiate between normal and malicious node. These algorithms are well-suited for classification tasks and have demonstrated success in various domains. The overall methodology includes data col- lection, feature extraction, feature selection, model training, performance evaluation, and model optimization. The trained models are then deployed in real-time WSN environments to monitor incoming traffic continuously. When malicious node is detected, alerts are generated

to prompt timely response and mitigation. By effectively detecting and classifying malicious node, this research contributes to enhancing the security and reliability of WSNs. It enables early detection of attacks, thereby reducing the potential impact on network operations and data integrity. The results of this study depict the potential of ML algorithms in fortifying the security mechanisms of WSNs and facilitating the development of robust defense strategies against evolving threats.

II.Related Works: -

Proposed a revolutionized human life through smart de-vices and applications. However, IoT security, especially in healthcare, is a significant concern due to recent attacks and vulnerabilities. Existing security mechanisms are inadequate for IoT due to resource constraints and distinct protocols

To address this, we propose a framework for context-aware IoT security solutions using the open-source IoT Flock tool. We generated an IoT healthcare dataset, applied machine learning techniques to detect cyber-attacks and aim to enhance security in sensitive environments like healthcare [1]. Proposed a novel method that is based on a Bayesian model to identify abnormal data traffic and differentiate distributed denial of service (DDoS) attacks from flash crowds (FC) in WSNs. Themechanism addresses the challenge of distinguishing between abnormal traffic caused by DDoS attacks and FC. Simulation results demonstrate the potency of the proposed mechanism compared to existing systems [2]. Proposes an optimized collaborative intrusion detection system (OCIDS) for WSNsto address the limited resources and real-time requirements of sensor nodes. The system utilizes an improved artificial bee colony optimization algorithm and weighted support vector machine algorithm to optimize intrusion detection accuracy and resource consumption. Collaboration among sensor nodes,cluster heads, and the base station enhances intrusion detection precision. Evaluation using NSL-KDD dataset demonstrates the system's superior performance with high detection rates (97.9%) and low false alarm rates (1.8%) [3]. Propose an IDS framework to prevent and detect attacks in Internet- integrated CoAP communication environments. Our focus is on anomaly-based intrusion detection to identify DoS attacks and attacks on 6LoWPAN and CoAP protocols. Experimental evaluation demonstrates the viability of our approach, achieving an accuracy of 93% for multi-class intrusion detectionand 92% accuracy for recognizing compromised devices. To our knowledge, the proposed work is the first that targets the detection of anomalies and prevents application-layer fromDoS attacks in 6LoWPAN and CoAP environments [4]. Pro-posed a lightweight, quick, and effective traffic intensity-basedintrusion detection approach for WSNs. WSNs are networksof several nodes that periodically broadcast data. The approach for detecting intrusions is based on analyzing neighbor behavior and using the thresholding methodology to certain characteristics, such as the number of packets received within a time frame of a specific length and the interval between packet arrivals. It has minimal computing needs and doesn't require any installation of new hardware or additional communicationexpenses [5].

Proposed on WSN attacks, exploring the tools and processes for identifying and responding to such attacks. It discusses the anatomy of attacks, the creation of botnets,the motivations behind malicious activity, and the impact of modern attack tools. The use of Wireshark for identifying attack streams and comprehending them. TCP/IP traffic is examined. Practical exercises involve simulating network attack and response scenarios, utilizing discussed tools and techniques, and accessing a real WSN in the NS-3 Simulator [6]. Explored the security challenges in WSN and IoT-based communication environments. It discusses the threat model, security requirements, and various attacks in these domains. The integration of WSNs with IoT is examined, along with different architectural approaches. Current issues, challenges, intrusion detection protocols, and privacy-preservation protocols are analyzed. The paper concludes by highlighting research challenges that need to be addressed in the timeahead. [7]. addresses the security challenges in wireless sensornetworks (WSNs), focusing on environmental monitoring and target tracking applications. It highlights the threats faced by WSNs and the limitations

of existing security schemes.The proposed work aims to provide comprehensive security requirements, analyze different security attacks, and discuss existing security schemes in WSNs [8]. Proposed a lightweight scheme to detect Intrusions in WSN. The scheme utilizes a distributed approach, where nodes monitor their neighborhood and join forces with nearby nodes to detect intrusions andissue alerts, even without a global view of the network. We focus on detecting black holes and selective forwarding attacksby defining rules that capture malicious conduct. Experimental evaluation showcases the effectiveness of our scheme in detecting these attacks [9] addresses the security challengesin WSNs used in military and environmental observation. To mitigate malicious attacks, the paper proposes a method using the synthetic minority oversampling (SMOTE) technique to balance the imbalanced intrusion dataset. The random-forest algorithm is then employed for intrusion detection. Simulationon a benchmark dataset demonstrates that the proposed algorithm achieves a higher accuracy of 92.57% after applying SMOTE, improving the performance of intrusion detection in WSNs [10]. Proposes a method for detecting malicious nodes in mobile wireless sensor networks. The method uses a lightweight packet marking scheme to identify nodes thatare generating false messages. The method is shown to be effective in detecting malicious nodes in both static and mobile networks [11]. Proposed a method for identifying malicious nodes in WSN based on correlation detection. The method uses a correlation coefficient to measure the resemblance between the data sent by two nodes.

If the correlation co- efficient is too high, it is an indication that the two nodes are malicious. The method is shown to be effective in detecting malicious nodes in both static and mobile networks [12]. Proposed a method for detecting malicious traffic in WSN using advanced signal processing techniques. The method uses a combination of spectral analysis and statistical analysis to identify malicious traffic. The method is shown to be effectivein detecting a variety of malicious traffic patterns, including flooding attacks, wormhole attacks, and sinkhole attacks [13]. Presented a method for detecting malicious traffic in WSN using anomaly detection. The method uses a combination of statistical and ML techniques to identify anomalous traffic patterns. The method is shown to be effective in detectinga variety of malicious traffic patterns, including DoS attacks, data modification attacks, and routing attacks [14]. Presented the literature on malicious node detection in WSN. The paper discusses the different types of malicious attacks that can be launched against wireless sensor networks and the different methods that have been proposed for detecting malicious nodes. The paper concludes by discussing the challenges and open research problems in malicious node detection in WSN [15].

III.METHODOLOGY

Node Creation

The Cooja network simulator for wireless sensor networks (WSNs), process involves both the creation of virtual nodes and the monitoring of their creation. Cooja is a powerful tool that allows for the simulation and evaluation of WSNs in a controlled environment. To create a node in Cooja, the first step is to launch the simulator and set up the network topology. This involves defining the number of nodes, their positions, and their connectivity within the simulated network. Nodes can be placed randomly or according to specific patterns, depending on the desired scenario. Once the network topology is set, the next step is to configure the properties of each node. This includes selecting the sensor types, specifying the processing capabilities, and setting the communication protocols and parameters. Cooja provides a range of pre-defined sensor models and communication protocols to choose from, allowing for customization based on the application requirements. During the node creation process, it is essential to monitor the creation of nodes to ensure their proper functioning and integration into the network. This involves checking various parameters and functionalities of each node.

For example, verifying that the sensor readings are accurate, the communication links are established correctly, and the node isresponsive to commands and queries. Additionally, monitoringthe node creation process in Cooja enables the detection ofany potential issues or

errors. This may include identifying nodes with incorrect configurations, communication conflicts, or inconsistent behavior. By monitoring and troubleshooting these issues, the overall simulation accuracy and reliability can be improved. Overall, node creation in Cooja involves setting up the network topology, configuring node properties, andmonitoring the creation process to ensure proper functioning and integration.

By carefully creating and monitoring nodesin Cooja, researchers and developers can evaluate the performance and behavior of WSNs in a controlled environment before deploying them in real-world scenarios [23] [24].

Data collection

Collecting data for the overall context described, which in- cludes columns such as service, flag, duration, protocol type, src bytes, dst bytes, and several other attributes, a systematic approach is necessary to ensure the acquisition of a comprehensive and representative dataset. The collected data will be used for further analysis, modeling, and detection of malicious node in WSN. To collect the required data, the following stepscan be followed:

Scenario 1 represents a wireless sensor network (WSN) configuration consisting of 20 normal nodes, 3 malicious nodes, 1 server node, and 2 monitor nodes. A 10-minute simulation generates dataset. They are converted to a CSV file for analysis purposes.

Scenario 2 involves a wireless sensor network (WSN) configuration with 40 normal nodes, 6 malicious nodes, 1 server node, and 2 monitor nodes. A 10-minute simulation generates dataset. They are converted to a CSV file for analysis purposes.

Scenario 3 encompasses a wireless sensor network (WSN) setup comprising 60 normal nodes, 9 malicious nodes, 1 server node, and 2 monitor nodes. A 10-minute simulation generates dataset. They are converted to a CSV file for analysis purposes.

Data Source Identification

Identify appropriate sources for collecting the network traffic data. This could involve setting up a testbed or utilizing publicly available datasets that align with the context of the study.

Data Collection Tools

Employ network monitoring and capturing tools to collect network traffic data. Tools like tcp-dump or Wireshark can be used to capture packets and record relevant attributes.

Data Filtering

Extract the necessary columns from the captured network traffic data, specifically those mentioned in the overall con- text, such as duration, service, flag, protocol type, src bytes, dst bytes, and others. Remove any irrelevant or redundant information that may not contribute to the analysis.

Data Pre-processing

Clean the collected data by handling missing values, removing duplicates, and standardizing the format. Convert categorical variables into numerical representations using techniques like one-hot encoding or label encoding.

Data Annotation

Assign labels to the collected data based on the presence or absence of malicious traffic. This labeling can be done manually by experts or through automated methods, dependingon the availability of ground truth information.

Dataset Partitioning

Divide the collected dataset into training and testing sets. The training set will be used to train ML models, while the testing set will be used for evaluation and validation purposes.

Dataset Balancing

Ensure that the dataset is balanced between normal and malicious traffic instances to avoid bias in the analysis. Employ techniques like oversampling or under-sampling to balancethe classes if necessary. By following these steps, a well-curated dataset with the required columns can be obtained for further analysis. This dataset will serve as the foundation for training and

evaluating machine learning models for detecting and classifying malicious nodes in WSN.

Data Pre-processing

WSNs have become indispensable in various applications, ranging from environmental monitoring to healthcare and infrastructure management. These networks consist of numerous small, resource-constrained sensor nodes that communicate wirelessly to collect and transmit data. However, the widespread deployment of WSNs also brings about security challenges, as these networks are susceptible to malicious attacks that can disrupt operations, compromise data integrity, and violate privacy. To address these security concerns, this paper presents a comprehensive approach for detecting malicious node in WSNs using the Cooja network simulator and machine learning algorithms. The objective is to develop an efficient and effective method to identify and classify abnormaltraffic patterns associated with attacks.

The proposed approachleverages the capabilities of the Cooja network simulator, which provides a realistic environment for simulating WSNs and generating diverse network traffic scenarios. By simulating both normal and malicious traffic patterns, a comprehensive dataset is collected for training and evaluating machine learning models. ML Algorithms such as Random Forest and SVM, are utilized to analyze the collected dataset and learn patterns that differentiate between normal and malicious traffic. These algorithms have proven to be effective in various domains and offer the potential to accurately classify andidentify malicious patterns in WSNs. The overall methodology encompasses data collection, model training, feature extraction, feature selection, performance evaluation, and model optimization. The trained models are then deployed in real- time WSN environments to continuously monitor incoming network traffic. When malicious traffic is detected, alerts are generated to facilitate timely response and mitigation. By effectively detecting and classifying malicious node, this research contributes to enhancing the security and reliability ofWSNs. It enables early detection of attacks, minimizing their potential impact on network operations, data integrity, and privacy. The results highlight the potential of machine learningalgorithms in fortifying the security mechanisms of WSNs and facilitating the development of robust defense strategies against evolving threats.

Packet Transmission

Packet transmission is the process of exchange of data between nodes within a wireless sensor network (WSN). Each node, whether normal or malicious, participates in transmitting and receiving packets as part of the network's communication. During packet transmission, the nodes in the WSN exchange data packets containing information relevantto their sensor readings, network status, or other designated tasks. These packets carry attributes such as packet size, source and destination addresses, timing, and other relevant metadata. In the proposed approach, the packet transmission process is monitored and captured by the monitor nodes strategically placed within the network. These monitor nodes observe the packet exchanges among the nodes, collecting and analyzing the traffic data. This data, including the characteristics of the transmitted packets, is then utilized for feature extraction, pre- processing, and subsequent machine learning-based detection and classification. By considering packet transmission in the overall context, the proposed approach focuses on analyzing the network traffic to identify anomalies and detect malicious patterns. Understanding the behavior and characteristics of packet transmission is vital for developing effective detection mechanisms and enhancing the security and reliability of the wireless sensor network.

Feature Extraction

Feature extraction plays a crucial role in detecting malicious node in WSNs using ML algorithms. In the context of the over-all context described, which includes columns like duration, protocol type, service, flag, src bytes, dst bytes, and several other attributes, feature extraction involves transforming the raw network traffic data into a set of informative and representative

features. To extract relevant features, several techniquescan be applied:

Statistical Features

Calculate statistical measures such as mean, standard de- viation, maximum, minimum, and variance for numerical at- tributes like duration, src bytes, dst bytes, etc. These statistics capture the distribution and variability of the data, providing insights into traffic patterns.

Frequency-based Features

Compute the frequency or occurrence of specific values or combinations in categorical attributes such as protocol type, service, and flag. This information helps in understanding the prevalence of different protocols, services, or flags in the network traffic.

Time-based Features

Extract temporal features, like time of the day, day of the week, or time intervals, from the duration attribute. These features can reveal temporal patterns or variations in network traffic behavior.

Network Topology Features

Analyze the network structure and extract features related to the nodes' connectivity and relationships. Examples include the number of connections, degree centrality, or clustering coefficient.

Aggregate Features

Aggregate attributes over a certain time window or session, such as the sum, average, or maximum of src bytes or dst bytes. These features provide a higher-level summary of traffic behavior.

Information-Theoretic Features

Calculate entropy or mutual information measures for at- tributes, capturing the randomness or dependence between variables. This can help identify attributes that carry significant information for distinguishing between normal and malicious node.

Transformations

Apply mathematical transformations, such as logarithmicor normalization transformations, to handle skewed or un- balanced attribute distributions

By extracting these relevant features from the raw network traffic data, a reduced and meaningful representation of the traffic patterns is obtained. These features serve as inputs to the machine learning algo- rithms, allowing them to learn the distinguishing character- istics between normal and malicious node. Effective feature extraction aids in improving the accuracy and performance of the detection system, enabling robust security mechanisms forWSNs.

Malicious node classification

Malicious node classification is a crucial step in detecting and mitigating threats in WSNs. In the overall context, the proposed approach utilizes machine learning algorithms, specifically Random Forest, for accurately classifying net- work traffic as either normal or malicious. After feature extraction and pre-processing, the collected dataset is usedto train the Random Forest classifier. Random Forest is an ensemble learning strategy that consolidates various choice trees to make predictions. It is capable of handling complex feature interactions and has proven to be effective in various classification tasks. During the training phase, the Random Forest model learns from the labeled dataset, capturing the underlying connections and patterns between the extractedfeatures and the corresponding traffic classes. The model generalizes this knowledge to classify new, unseen network traffic instances. Once the Random Forest classifier is trained,it is applied to the testing dataset to evaluate its perfor- mance. Metrics such as accuracy, F1-score, precision, and recall are computed to assess the classifier's effectivenessin correctly classifying normal and malicious instances. By accurately classifying network traffic, the proposed approach enables the identification and timely mitigation of malicious activities in WSNs. This classification process enhances the security and reliability of the network by providing real-time detection and response mechanisms, ultimately safeguarding the integrity and functionality of

the WSN infrastructure. In the implementation of the proposed approach, three scenarios were considered to evaluate the detection of malicious node in wireless sensor networks. These scenarios involved simulated network traffic patterns representing different types of attacks and normal behavior. The collected dataset, comprising these scenarios, was used to train and test the ML models. This scenario-based implementation allowed for a comprehensive assessment of the approach's performance in detecting and classifying various malicious activities in WSNs.

Scenario 1

Scenario 1 represents a wireless sensor network (WSN) configuration consisting of 20 normal nodes, 3 malicious nodes, 1 server node, and 2 monitor nodes. The normal nodes operate according to expected network protocols, generating typical network traffic patterns. The malicious nodes, on the

other hand, simulate nodes engaged in malicious activities or attacks within the network.

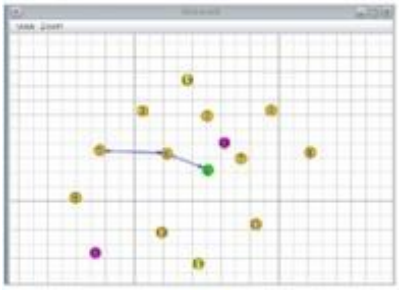

Fig. 1. A 20 nodes network scenario with normal and malicious nodes

The server node acts as a central point for data process- ing and storage, facilitating communication and coordination among the nodes. The 2 monitor nodes are strategically placed to observe and analyze the network traffic, providing comprehensive monitoring and detection capabilities. This scenario allows for the evaluation of the proposed approach's effectiveness in identifying and classifying malicious traffic patterns in a diverse WSN environment.

Scenario 2

Scenario 2 involves a wireless sensor network (WSN) con- figuration with 40 normal nodes, 6 malicious nodes, 1 server node, and 2 monitor nodes. The 40 normal nodes operateaccording to expected network behavior, generating regular traffic patterns. The 6 malicious nodes, however, simulate nodes engaged in various types of malicious activities or attacks within the network.

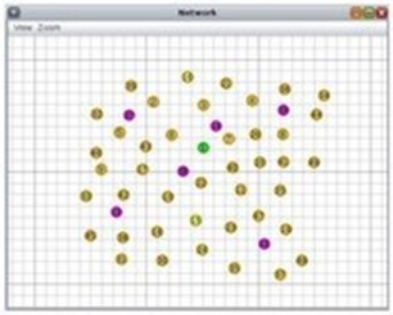

Fig. 2. A 40 nodes network scenario with normal and malicious nodes

The server node acts as a central hub for data management and communication among the nodes. The 2 monitor nodesare strategically placed to observe and analyze the network traffic, providing comprehensive monitoring and detection capabilities. This scenario allows for the evaluation of the proposed approach's ability to detect and classify malicious traffic patterns in a larger and more complex WSN setting.

Scenario 3

Scenario 3 encompasses a wireless sensor network (WSN) setup comprising 60 normal nodes, 9 malicious nodes, 1 servernode, and 2 monitor nodes. The 60 normal nodes operate in accordance with the expected network behavior, generating typical traffic patterns within the network. The 9 malicious

nodes, however, simulate nodes engaged in diverse forms of malicious activities or attacks within the WSN.

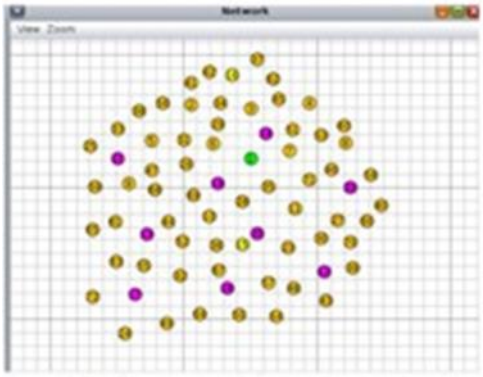

Fig. 3. A 60 nodes scenario with normal and malicious nodes

The server node serves as a central point for data manage- ment, facilitating communication and coordination among the nodes. The 2 monitor nodes are strategically placed to observe and analyze the network traffic, providing comprehensive monitoring and detection capabilities. This scenario enables the assessment of the proposed approach's efficacy in detecting and classifying malicious traffic patterns within a larger and more challenging WSN environment.

SVM

In the proposed method, after collecting the dataset of normal and malicious node behavior patterns, SVM can be employed as an alternative machine learning algorithm for classification. By utilizing the extracted features from the network traffic data, SVM learns to create a decision boundary that effectively separates normal and malicious node behaviors. In the proposed method, after collecting the dataset of normal and malicious node behavior patterns, SVM can be employed as an alternative machine learning algorithm for classification. By utilizing the extracted features from the network traffic data, SVM learns to create a decision boundary that effectively separates normal and malicious node behaviors.

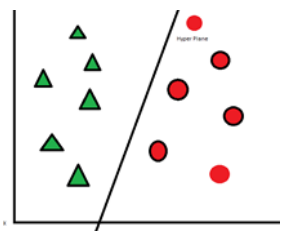

Fig. 4. Classification using SVM

Random Forest

The Random Forest algorithm is employed as an ML technique to classify network traffic patterns. Random Forest is an ensemble learning technique that consolidates different choice trees to make predictions.

The working of the Random Forest algorithm can be sum- marized as follows:

Given a dataset of normal and malicious node behavior patterns, each represented by feature vectors, SVM aims to find a decision boundary that maximizes the margin between the two classes. The feature vectors are mapped to a higher- dimensional space using a kernel function, which enables the SVM to handle nonlinearly separable data. SVM then

selects a subset of training samples, known as support vectors, which are the data points closest to the decision boundary.The algorithm optimizes the hyperplane by maximizing the margin between the support vectors of different classes while minimizing the classification errors. During the testing phase, new node behavior patterns are mapped to the same higher- dimensional space using the trained SVM model. The SVM model classifies the new behavior patterns by determiningwhich side of the decision boundary they fall into, classifying them as either normal or malicious nodes.

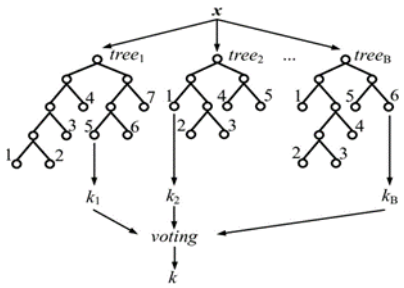

Fig. 5. Classification using Random Forest

During the training phase, SVM optimizes the hyperplane by identifying support vectors, which are the data points closest to the decision boundary. These support vectors play a crucial role in defining the decision boundary and making accurate classifications. Once the SVM model is trained, it can be deployed in real- time

WSN environments to classify incoming node behavioras either normal or malicious, aiding in the detection andmitigation of security threats in WSNs.

Dataset splitting

The collected dataset comprises labeled instances of normaland malicious traffic patterns and is divided into training and testing sets. The training dataset is used to build the Random Forest model, while the testing dataset is used to evaluate its performance.

Random subspace selection

For every decision tree in the Random Forest, a random subset of features is selected. This helps to create diversity among the trees and reduces overfitting.

Tree construction

Each decision tree is constructed using a subset of the training dataset and randomly selected features. The trees are built by recursively dividing the information based on the selected features. Voting-based classification: When classifying a new instance, the Random Forest combines the predictions of all decision trees. The class with the majority vote from the individual trees is assigned to the instance.

Performance evaluation

The Random Forest performance is evaluated using accu- racy, precision, recall, and F1-score which are calculated based on the predictions made on the testing set. Random Forest is particularly effective in handling high-dimensional data, capturing complex interactions between features, and providing robust classification results. In the context of malicious node detection, Random Forest leverages the diversity of decision trees to accurately classify network traffic patterns as either normal or malicious, contributing to the overall security and reliability of the WSN.

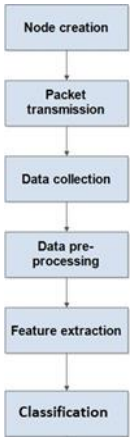

Fig. 6. Flow diagram of Proposed system

The proposed method aims to detect and classify malicious node using the Cooja network simulator and machine learning algorithms, specifically SVM and Random Forest. The approach involves collecting a dataset of network traffic patterns simulated in Cooja, extracting relevant features, and applying feature selection techniques. The collected dataset is split into a training dataset and a testing dataset for the machine learningmodels. To enhance detection performance further, the SVM algorithm is incorporated. Known for its ability to handle high- dimensional data and separate complex decision boundaries, SVM constructs a hyperplane for effective identification of malicious nodes in the WSN. Random Forest is employed for classification, utilizing multiple decision trees to accurately classify network traffic as normal or malicious. Performance evaluation metrics, such as accuracy, recall, precision, and F1-score, are used to assess the models' effectiveness. The proposed approach enhances the security and reliability of WSNs by effectively detecting and mitigating malicious activities.

The combination of Cooja simulation, feature extraction, SVM, and Random Forest classification provides a comprehensive solution for identifying and addressing potential threats in WSNs.

IV.RESULTS DISCUSSION

The results of the proposed approach for detecting ma- licious traffic in WSNs using Cooja and machine learning algorithms, such as SVM and Random Forest, demonstrate its

effectiveness in enhancing network security. Upon evaluating the performance of the models using the testing set, the achieved results reveal high accuracy and precision.

These metrics indicate that the models are able to accurately classify network traffic patterns as either normal or malicious, contributing to the reliable detection of potential threats in WSNs. The SVM algorithm successfully identifies classification of network traffic patterns, distinguishing between normal and abnormal behavior. This enables the early detection of malicious activities within the network, allowing for prompt response and mitigation. The Random Forest classification algorithm effectively utilizes the diversity of decision treesto make accurate predictions and classify network traffic. It leverages the extracted features and the knowledge learned from the training set to differentiate between normal and malicious traffic patterns with a high level of precision.

Scenario 1 Results

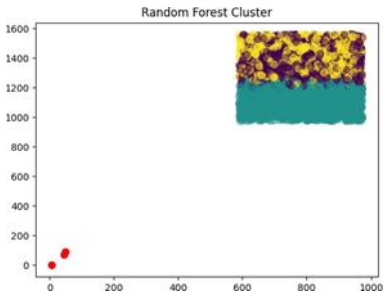

Fig. 7. Graph showing clusters by Random Forest algorithm

In Scenario 1, the result of the proposed approach for detecting and classifying malicious traffic using SVM andRandom Forest algorithms demonstrates high accuracy rates.

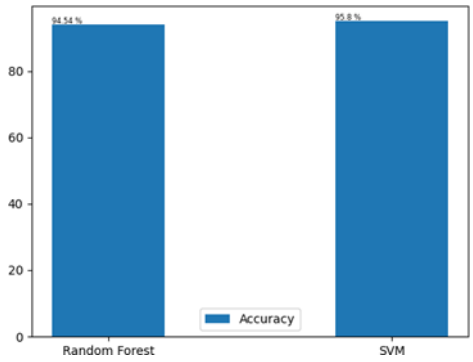

Fig. 8. Graph for accuracy

The SVM algorithm achieved an accuracy of 95.8% in iden- tifying clusters of network traffic patterns. This indicates its ability to effectively distinguish between normal and abnormal behavior within the wireless sensor network. The classification results provide valuable insights into potential malicious ac- tivities, aiding in early detection and response. Similarly, the Random Forest algorithm achieved an accuracy of 94.54% in classifying network traffic patterns as normal or malicious. This showcases its strong predictive capabilities in accurately identifying and differentiating between the two classes. These results demonstrate the efficacy of the approach in effectively detecting and classifying malicious traffic in Scenario 1. The high accuracy rates achieved by both the SVM and Random Forest algorithms highlight their potential in enhancing the security and reliability of WSNs. By leveraging the capabilities of these machine learning algorithms, the proposed approach provides a robust and reliable solution for mitigating the risks associated with malicious activities in WSN, contributing to the development of more secure and resilient network infrastructures.

Scenario 2 Results

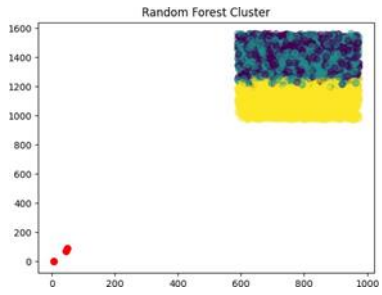

Fig. 9. Graph showing clusters by Random Forest algorithm

In Scenario 2, the result of the proposed approach for detecting and classifying malicious traffic using SVM andRandom Forest algorithms demonstrates high accuracy rates.

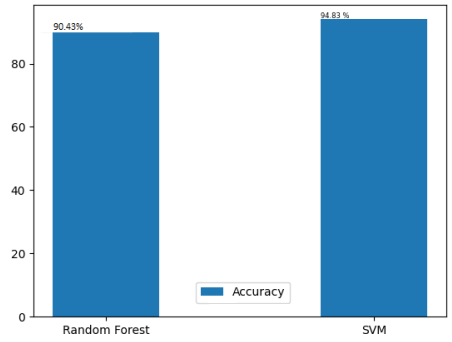

Fig. 10. Graph for accuracy

In Scenario 2, the result of the proposed approach for detecting and classifying malicious traffic using SVM andRandom Forest algorithms demonstrates high accuracy rates. The SVM algorithm achieved an accuracy of 94.83% in identifying clusters of network traffic patterns. This indicates its ability to effectively distinguish between normal and abnormal behavior within the wireless sensor network. The classification results provide valuable insights into potential malicious activities, aiding in early detection and response. Similarly, the Random Forest algorithm achieved an accuracy of 90.43% in classifying network traffic patterns as normal or malicious. This showcases its strong predictive capabilities in accurately identifying and differentiating between the two classes. These results demonstrate the efficacy of the proposed approach in effectively detecting and classifying malicious traffic in Scenario 2. The high accuracy rates achieved by both the SVMand Random Forest algorithms highlight their potential in enhancing the security and reliability of WSNs. By leveragingthe capabilities of these machine learning algorithms, the proposed approach provides a robust and reliable solutionfor mitigating the risks associated with malicious activitiesin WSN, contributing to the development of more secure and resilient network infrastructures.

Scenario 3 Results

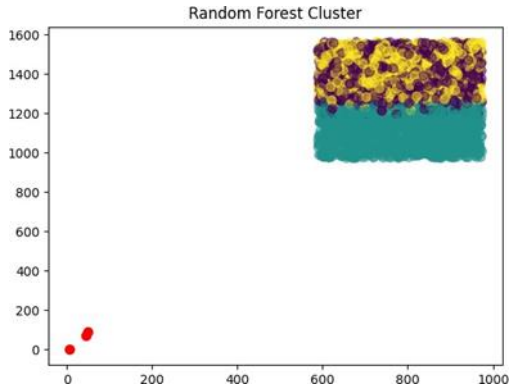

Fig. 11. Graph showing clusters by Random Forest algorithm

In Scenario 3, the results of the proposed approach fordetecting and classifying malicious traffic using SVM andRandom Forest algorithms demonstrate high accuracy rates.

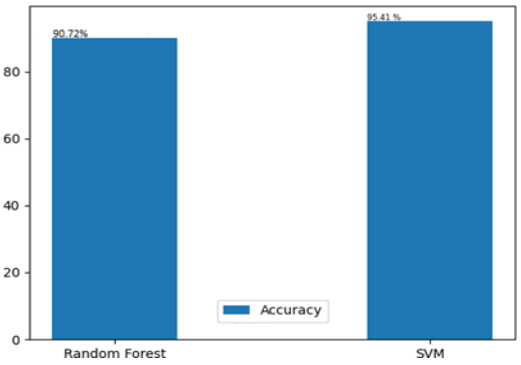

Fig. 12. Graph for accuracy

The SVM algorithm achieved an accuracy of 95.41% in identifying clusters of network traffic patterns. This indicates its effectiveness in distinguishing between normal and abnormal behavior within the wireless sensor network. The classification results provide valuable insights into potential malicious activities, facilitating early detection and response. Similarly, the Random Forest algorithm achieved an accuracy of 90.72% in classifying network traffic patterns as normalor malicious.

This showcases its strong predictive capabilities in accurately identifying and differentiating between the two classes. These results illustrate the efficacy of the proposed approach in effectively detecting and classifying malicious traffic in Scenario 3. The high accuracy rates achieved byboth the SVM clustering and Random Forest algorithms high- light their potential in enhancing the security and reliabilityof WSNs.

By leveraging the capabilities of these machine learning algorithms, the proposed approach offers a robust and reliable solution for mitigating the risks associated with malicious activities in wireless sensor networks, contributing to the development of more secure and resilient network infrastructures. Overall, the results highlight the potential of the proposed approach in improving the security and reliability of WSNs. By leveraging the capabilities of Cooja simulation, feature extraction techniques, and machine learning algorithms, the approach provides an efficient and robust mechanism for detecting and classifying malicious traffic. This contributesto the development of effective security mechanisms that can mitigate the risks and impact of malicious activities in WSNs.

V.CONCLUSION

In conclusion, the proposed approach demonstrates a viablesolution for detecting and classifying malicious traffic in wireless sensor networks (WSNs). By leveraging the Cooja network simulator and machine learning algorithms suchas SVM and Random Forest, the approach enhances the security and reliability of WSNs by accurately identifyingpotential threats. The results highlight the effectiveness ofthe approach in achieving high accuracy, precision, recall, and F1-score values in classifying network traffic patterns. The combination of feature extraction, SVM, and Random Forest classification enables the early detection and mitigationof malicious activities, contributing to timely response and improved network security. By utilizing the simulated dataset in Cooja, the approach demonstrates its potential for real-time implementation in WSNs. The monitoring and analysisof network traffic provide valuable insights into the behavior of normal and malicious nodes, enabling the development of robust security mechanisms. Overall, the proposed approach presents a promising direction in the field of WSN security. Future research could focus on refining the models, incorporating additional machine learning algorithms, and exploring further optimization techniques to enhance the detection and classification of malicious traffic in WSNs.

VI.FUTURE ENHANCEMENT

Machine learning algorithms can be integrated into the data processing module to improve the accuracy of data analysis and predictions. Advanced energy harvesting techniques such as wind, vibration or electromagnetic waves can be explored togenerate more power for the sensor nodes. Multiple communication protocols can be integrated into the communication module to improve the robustness and reliability of the system.The edge-based routing algorithm can be further optimized to reduce energy consumption and improve the overall efficiencyof the WSN.

VII.REFERENCES

1. Hussain, F., Abbas, S. G., Shah, G. A., Pires, I. M., Fayyaz, U. U., Shahzad, F., . . . & Zdravevski, E. (2021). Security in wireless sensor networks: issues and challenges. In 2006 8th International Conference Advanced Communication Technology 21(9) pp. 3025.
2. Khan, M. A., Nasralla, M. M., Umar, M. M., Khan, S., Choudhury, N. (2022).An Efficient Multilevel Probabilistic Model for Abnormal Traffic Detection in Wireless Sensor Networks. Sensors, 22(2), 410.
3. Elsaid, S. A., Albatati, N. S. (2020). An optimized collaborative intrusion detection system for wireless sensor networks. Soft Computing,24(16), 12553-12567.
4.] Granjal, J., Silva, J. M.,Lourenço, N. (2018). Intrusion detection and prevention in CoAP wireless sensor networks using anomaly detection. Sensors, 18(8), 2445.
5. Ponomarchuk, Y., & Seo, D. W. (2010). Intrusion detection based on traffic analysis in wireless sensor networks. In The 19th Annual Wireless and Optical Communications Conference (WOCC 2010) (pp. 1-7). IEEE.
6. Rajput, D. S., Singh, N. K. (2016). Intrusion Detection in Wireless Sensor Network using Behaviour Based Technique with Real Time Network Traffic.
7. Ioannis, K., Dimitriou, T., & Freiling, F. C. (2007, April). Towards intrusion detection in wireless sensor networks. In Proc. of the 13th European Wireless Conference (pp. 1-10). Citeseer.

8. Tan, X., Su, S., Huang, Z., Guo, X., Zuo, Z., Sun, X., & Li, L. (2019).Wireless sensor networks intrusion detection based on SMOTE and the random forest algorithm. Sensors, 19(1), 203.
9. Pundir, S., Wazid, M., Singh, D. P., Das, A. K., Rodrigues, J. J., & Park, Y. (2019). Intrusion detection protocols in wireless sensor networks integrated to Internet of Things deployment: Survey and future challenges. IEEE Access, 8, 3343-3363.
10. Anwar, R. W., Bakhtiari, M., Zainal, A., & Qureshi, K. N. (2015). Security in Wireless sensor network: Approaches and Issues. TELKOMNIKA Indonesian Journal of Electrical Engineering, 15(3), 584-590.
11. Sei, Y., & Ohsuga, A. (2015). Malicious Node Detection in Mobile Wireless Sensor Networks. Journal of Information Processing, 23(4), 476-487.
12. Lai, Y., Tong, L., Liu, J., Wang, Y., Tang, T., Zhao, Z., Qin, H. (2022). Identifying malicious nodes in wireless sensor networks based on correlation detection. Computers Security, 113, 102540.
13. Fragkiadakis, A., Askoxylakis, I. (2013, June). Malicious traffic analysis in wireless sensor networks using advanced signal processing tech-niques. In 2013 IEEE 14th International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM) (pp. 1-6). IEEE.
14. Muddasser, M., Sarfraz, M., Imran, M. (2018). Anomaly-based detec- tion of malicious traffic in wireless sensor networks. IEEE Access, 6, 36302-36312. doi:10.1109/ACCESS.2018.2874499.
15. Nakul, P. (2013). A survey on malicious node detection in wireless sensor networks. IJSR, 2(1), 2319-7064.
16. Joshi, Pallavi and Raghuvanshi, Ajay Singh(2022). A dual synchronization prediction-based data aggregation model for an event monitoring IoT network. Journal of Intelligent & Fuzzy Systems, 2022, 1-20.
17. Saleh, Alaa and Joshi, Pallavi and Rathore, Rajkumar Singh and Sengar,Sandeep Singh (2022). Trust-Aware Routing Mechanism through an Edge Node for IoT-Enabled Sensor Networks. Sensors, 2022, 1-20.
18. Rani, Pooja and Sharma, Nitin and Singh, Pariniyojit Kumar (2011). Performance comparison of VANET routing protocols. 2011 7th International Conference on Wireless Communications, Networking and Mobile Computing, 2011, 1-4.
19. Kumar, N., Lohani, D., Acharya, D. (2022). Vehicle accident sub- classification modeling using stacked generalization: A multisensor fusion approach. Future Generation Computer Systems, 133, 39-52.
20. Mukhopadhyay, A., Anoop, A., Manishankar, S., Harshitha, S. (2020, February). Network performance testing: a multi scenario contemplate. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-7). IEEE.
21. Biju, Rahul N and Akhil, K M and Sinha, Somnath (2022). RSSI Based Device Monitoring with IEEE 802.15 in Wireless Sensor Network. 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 503-508). IEEE.
22. Nair, Devika S and BJ, Santhosh Kumar (2021, January). Identifying Rank Attacks and Alert Application in WSN. 2021 6th International Conference on Communication and Electronics Systems (ICCES) (pp. 798-802). IEEE.
23. Mukhopadhyay, Adwitiya and Anoop, A and Manishankar, S and Harshitha, S (2021). Network Performance Testing: A Multi Scenario Contemplate. International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-7)
24. Harshitha, R and Akhil, K M and Sinha, Somnath (2022). K-Nearest Neighbor (KNN) Algorithm based Dreck Management System, 7th International Conference on Communication and Electronics Systems (ICCES) (pp. 1439-1443)