**International Innovative Research Journal of Engineering and Technology**
Vol. 10, Issue. 1, Sep 2024, pp. 9~17
ISSN: 2456-1983, DOI: 10.32595/iirjet.org/v10i1.20224.197


# Enhancing RPL Security in the Internet of Things for Preventive Network Threats with Machine Learning and NANTAR

**[1]Dr. K. Baskaran, [2]Dr.R.Madhubala**
[1]Associate professor, Department of Biomedical Engineering, Chennai institute of technology, Chennai, India.
[2] Lecturer, Information Technology Department, University of Technology and Applied Sciences – Shinas, Oman.

**Article Info**

***Article history:***

Received Jun 8, 2024
Revised Jul 19, 2024
Accepted Aug 30, 2024

***Keywords:***

Internet of Things
Routing Protocol for Low-Power and Lossy Networks(RPL)
Machine learning
Adaptive routing
Senor Network

**ABSTRACT**

"Internet of Things" refers to a collection of physical devices that are connected to the Internet and can communicate with one another. Rank, Sinkhole, and Wormhole attacks can seriously impair the performance of the Routing Protocol for Low-Power and Lossy Networks (RPL), which is essential for effective data transfer in Internet of Things networks. Due to its lightweight core, RPL cannot support resource-intensive and computationally intensive security implementation techniques. As a result, security attacks, which may be roughly divided into RPL-specific and sensor-network-inherited assaults, can affect both IoT and RPL. They take advantage of RPL resources and components, such as maintenance methods, routing settings, administrative messages, and network sensor components. The research presented here suggests an innovative machine learning-based method to increase RPL's security. We present a Novel Algorithm for Network Traffic Analysis and Response (NANTAR). This innovative approach that uses AI-based techniques to optimize the security and efficacy of RPL routing, in conjunction with a reinforcement learning module for dynamic and adaptive routing. This framework significantly improves packet false positive rate, lowers latency, and boosts network throughput while maintaining minimal jitter.

It effectively secures the RPL protocol in IoT-enabled wireless sensor networks by achieving a high detection rate, few false positives, and quick reaction to security problems.

***Corresponding Author:***

Dr. K. Baskaran,
Associate professor, Department of Biomedical Engineering,
Chennai institute of technology.
Email: baskarank@citchennai.net

## 1. INTRODUCTION

The Internet of things (IoT) has been regarded as the most important technological revolution of this period due to its frightening rate of expansion in application across all fields [1]. Sensor specialist Kevin Ashton coined the term "IoT" in 1992 to refer to the network that links physical objects to the Internet. IoT enhances productivity, ease of access, safety, and quality of life, which helps a wide range of industries and applications. It boosts their business and is utilized in several sectors. By increasing system productivity and efficacy in terms of time savings and function automation, as well as by adding intelligence to objects, IoT is reportedly on its way to enabling the modern world [2]. Also, IoT devices or objects may establish a bidirectional connection to exchange data via the Internet. In addition, the vast volumes of private information transmitted online will be vulnerable to security breaches from multiple angles, compromising IoT-enabled systems. The routing protocol, which creates and maintains paths between the network's devices, is one of the essential parts of IoT networks. Wireless sensors are useful in this situation. The routing

protocol determines how data packets are routed via intermediary devices from the source device to the target device. Due to its significant vulnerability to traffic and network threats, the network layer of these systems is a particular focus of analysis and research [3].

The Internet of Things uses a predefined IPv6 protocol called RPL. RPL is designed to function in limited resource environments, such as on devices with low capacity, storage, and long-term power. It can be applied to multiple paths schemas to improve energy efficiency, end-to-end packet delivery, and packet delivery rate [4]. The lack of a routing protocol that could satisfy the requirements of lossy, low-power, and resource-constrained Internet of Things networks caused the development of the RPL protocol. The distance vector routing theory generates a directed acyclic graph (DAG) or destination-oriented DAG (DODAG) to perform the routing tasks. RPL uses a metric called rank to calculate how different the devices in the DODAG are from one another. The five main control messages responsible for establishing a communication array are the DODAG information solicitation (DIS), DODAG information object (DIO), destination advertisement object (DAO), DAO acknowledgment (DAO-Ack), and consistency check (CC) messages [5]. This method is based on nodes. The nodes that act as gateways to the Internet are known as root nodes, and a number of parent nodes either directly or indirectly connect the other nodes in the network to them.

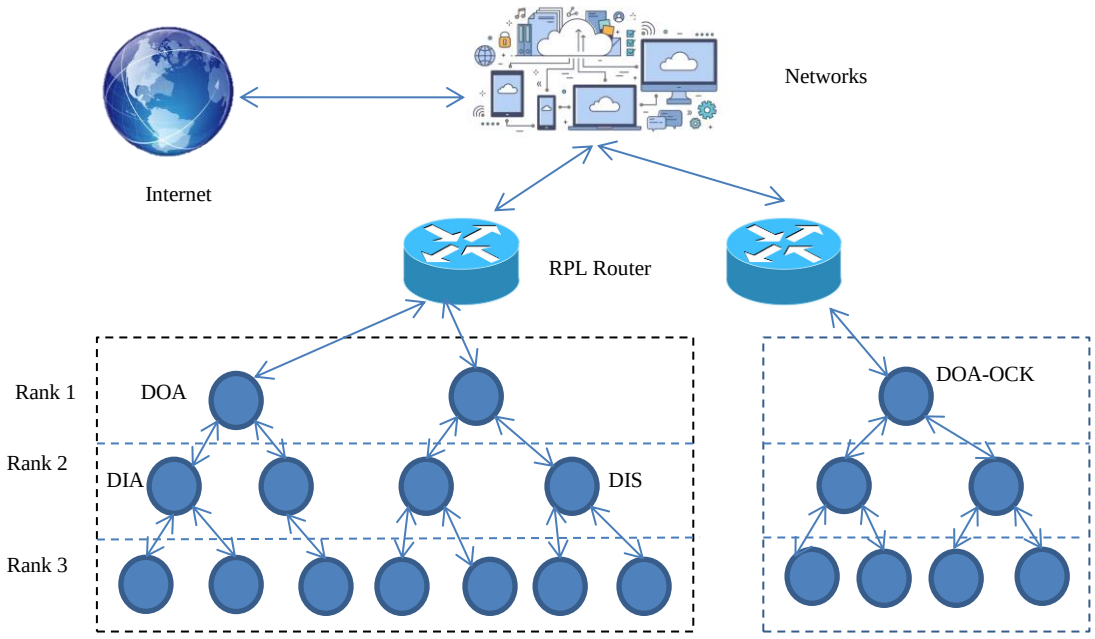


Figure 1. Architecture of DODAG

The above figure 1 represents the functions of the full architecture of DODAG [6]. This diagram illustrates how RPL maintains network routes and route information. RPL offers mechanisms for loop avoidance, local repair, and mobility support in addition to supporting many operating modes, including storing and non-storing modes [7].

Numerous attacks that seek to damage RPL's control messages or rank value place it in danger. For example, when a malicious device publishes a lower rank number than the actual rank number to boost traffic from other devices, this is known as a rank attack. Effective and efficient methods must be used to defend RPL networks from attacks [8]. A Novel Algorithm for Network Traffic Analysis and Response (NANTAR) is proposed in this paper to enhance RPL security in IoT platforms [9]. One of the main contributions of this article was the application of the NANTAR method, which leverages the Random Forest methodology for network traffic analysis and complex packet classification. NANTAR has shown a 29% throughput benefit 40–50% latency reduction, and more effective resource utilization compared to normal RPL. Major improvements in crucial operational KPIs are indicated by NANTAR's adaptability and seamless integration with different IoT platforms [10].

## 2. RELATED WORKS

Zhara et al [11] proposed a Rank and Wormhole Attack Detection Model for RPL-Based Internet of Things Using Machine Learning. The routing protocol for low-power and lossy networks (RPL) is used by the Internet of Things (IoT) to facilitate data exchange amongst its devices. Due to its lightweight core, RPL

is unable to support resource-intensive and computationally demanding security implementation techniques. To achieve multiclass classification, the model was benchmarked against the LIoTN-RPL dataset. With an average multiclass categorization of RA and WHA, the results were encouraging and superior to the prior model. Accuracy, precision, and recall for the multiclass LIoTN-RPL dataset were 99.7%, 99%, and 99.7%, respectively. At this point, it includes information on WHA, RA, and benign traffic.

Neerugatti et al [12] introduced a machine learning-based technique for detection of rank attack in RPL-based Internet of Things Networks. The Internet of Things has numerous applications in practically every industry. Its problems include privacy and security, robustness, weak points, low power, etc. Security is one of these main challenges. In networks based on the Internet of Things, RPL is a routing protocol that creates a path specifically for the limited nodes. The true positive and false positive rates of the suggested detection mechanism were computed using 30 motes in the Cooja simulation of the suggested technique. The malicious node was identified by comparing the computed rank with the node's initial rank. It was shown that the node identification rate in a network with 30 nodes is high and that the proposed method performed well in terms of delivery rate and latency.

Al-Amiedy et al [13] designed a systematic literature review on the machine and deep learning approaches for detecting attacks in RPL-Based 6LoWPAN of the Internet of Things. The environment is difficult to secure due to the limited resources of 6LoWPAN nodes, making it susceptible to threats and security attacks. Methods such as deep learning (DL) and machine learning (ML) have demonstrated promise as effective and successful ways to detect anomalous behavior in 6LoWPANs based on RPL. This study conducted a thorough and systematic literature review of the existing ML, DL, and combination algorithms to detect attacks in RPL-based 6LoWPAN networks. The fact that most researchers computed accuracy metrics for the measurements in their study shows how important they are to performance.

Seyfollahi et al [14] proposed A Review of Intrusion Detection Systems in RPL Routing Protocol Based on Machine Learning for Internet of Things Applications. At the network layer, the RPL protocol may generate routing and distribute routing data among nodes in the best possible way. RPL is a high-throughput, low-power IPv6 routing protocol that makes use of distance vectors. Numerous kinds of security threats can target the RPL protocol. Additionally, it uses two security modes to encrypt data packets. An invader could be a malfunctioning or incorrectly configured node whose actions impair network performance. An overview of IoT intrusion research initiatives is provided in this publication. The literature examined IoT or IoT intrusion detection methods that may be incorporated into an intrusion detection system.

Attique et al [15] introduced An Intrusion Detection System Using Fog and Deep Learning for RPL-Based Resource-Constrained Smart Industries. IoT is vulnerable to a wide variety of potential security risks due to its easily attainable nature. This work focuses on intrusion detection in RPL-based resource-constrained smart companies. This paper suggests a Fog-assisted DL-enabled intrusion detection framework (Cu-DNNGRU) to examine a variety of potential security flaws in smart businesses. The proposed framework is then compared with several well-known DL classifiers, such as Cu-LSTMDNN, Cu-BLSTM, and Cu-GRU, for a comprehensive performance analysis. The systematic simulation results, which demonstrate 99.39% accuracy, 99.09% precision, 98.89% recall, and 99.21% F1 score, validate the efficacy of the proposed model. Training the suggested model on a variety of datasets, seeks to enhance its detection skills going forward.

## 3. PROPOSED METHODOLOGY

This section includes a brief summary of each study phase and a thorough explanation of the technique used in the proposed study. The main components of this proposed system are NANTAR (Novel Algorithm for Network Traffic Analysis and Response), model training and selection, data collection, and data preparation. The goal of the approach is to classify network traffic using machine learning and to recognize the wormhole, rank, and DIS—three prevalent and dangerous attack types in IoT networks. Machine learning (ML) algorithms are proposed to develop a defense against wormhole and rank attacks.

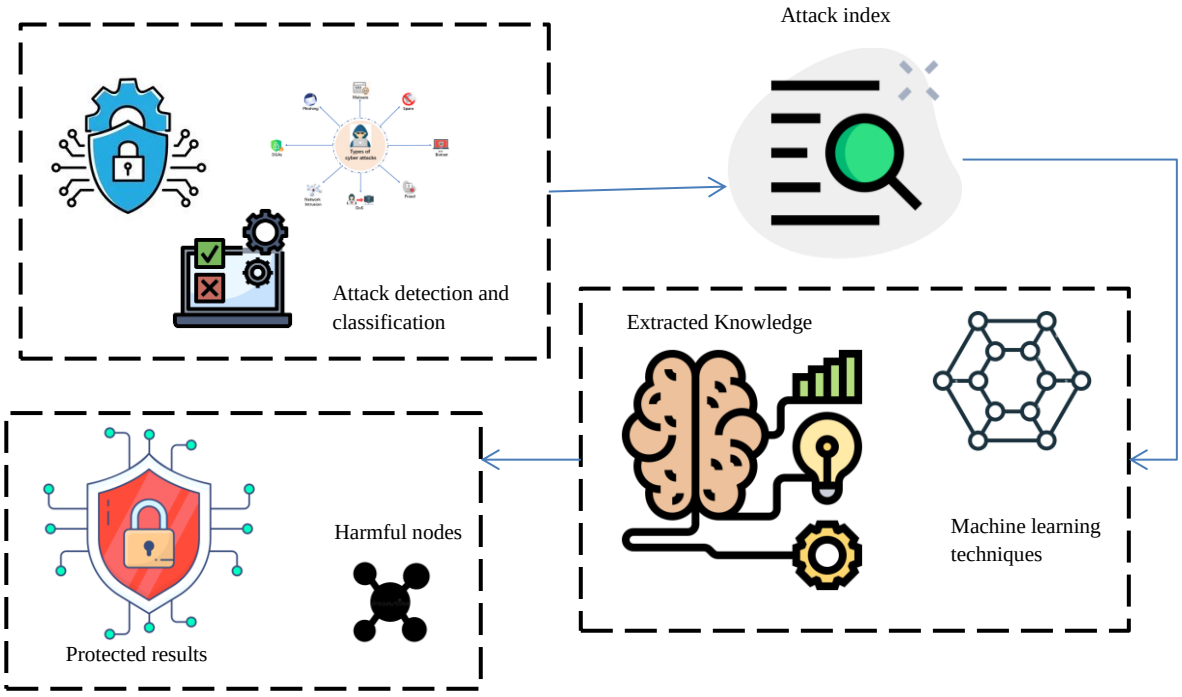


Figure 2. Harmful attacks Mitigation

In Figure 2 the attacker nodes are subsequently isolated and the RPL routing information is updated according to the classification results. For the following actions, the protected result is sent to the NANTAR model. The following are the methodology's specifics:

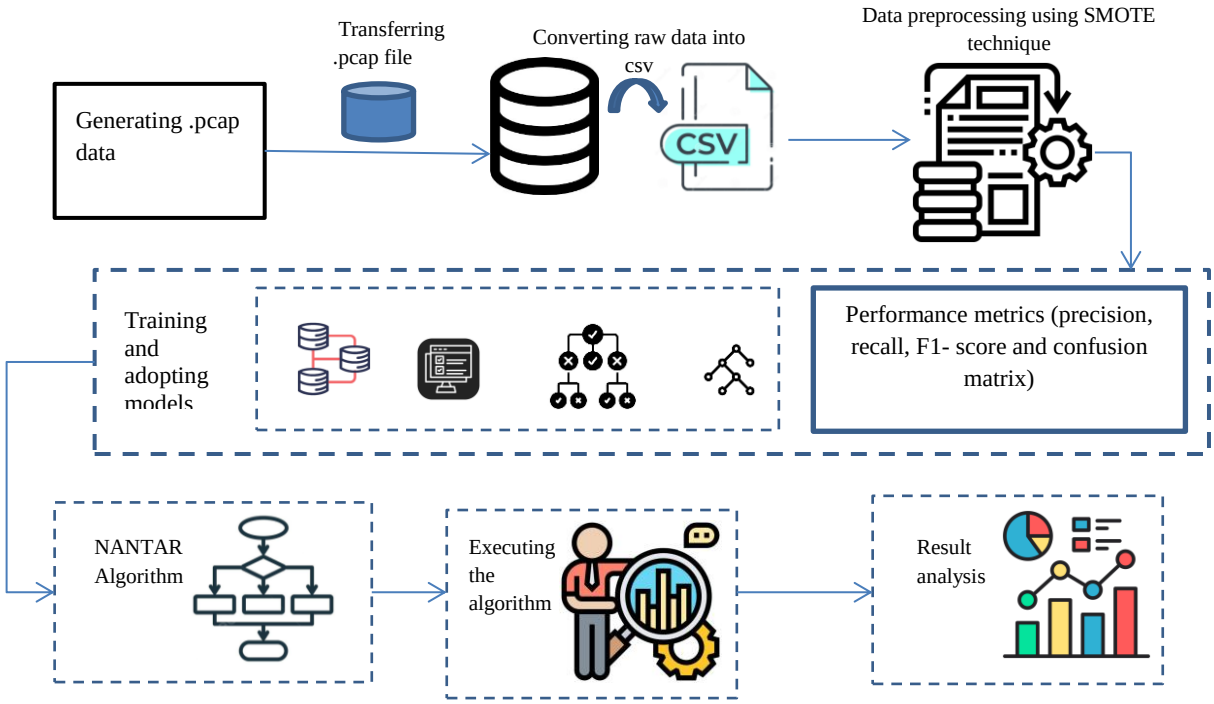


Figure 3. Architecture of proposed NANTAR system

In above figure 3 the architecture of the proposed system is given. The first concept is data collection. The data was produced by using the Cooja simulator, a network simulator for low-power wireless

devices, to simulate various network topologies and scenarios with variable numbers of nodes, traffic patterns, and attack kinds. Several procedures, including eliminating duplicates, outliers, missing values, and normalizing numerical features, were carried out in the data preprocessing section. Since the normal set is substantially larger than the malicious set, the dataset was balanced by using the SMOTE approach. The following formula can be used to create synthetic data using the SOMTE technique:

$$X_i = Y_i + \lambda \times (X_{mm} + Y_i) \quad (1)$$

Synthetic data are generated in Equation 1 for comparison with the real data. Where $X_{mm}$ refers to the nearest neighbor class and $Y_i$ is the minority class. Various machine-learning techniques were implemented and assessed in the third section. These models are supervised learning algorithms that are able to predict the class of new data by learning from labeled data. Accuracy, precision, recall, F1-score, and confusion matrix are the performance measurements we apply. The first evaluation metric used to assess the model's performance is accuracy. A greater number of algorithms were compared. It determines how many of the model's predictions were accurate out of all of them.

$$\frac{(RP + RN)}{(RP + RN + FP + FN)} \quad (2)$$

The accuracy is shown in equation (2), where RP, RN, FP, and FN stand for real positive, real negative, false positive, and false negative, in that order.

$$\frac{(RP)}{(RP + FP)} \quad (3)$$

Precision is the second statistic used to assess the performance of the model. It uses the confusion matrix's parameters to determine each class's accuracy. The precision can be found in equation (3).

$$\frac{RP}{(RP + FN)} \quad (4)$$

The ratio between the number of attacks the system detects and the total number of attacks in the dataset is called recall, or detection rate. The detection rate is computed using equation (4). We use random forest algorithms, which are the best when compared to other machine learning techniques. This type of collaborative learning incorporates numerous decision trees and uses majority voting to reach a conclusion. And its accuracy scores are great. Additionally, it does a decent job of differentiating between the classes. The algorithm for this system is then created.

### 3.1. NANTAR Algorithm development

The Random Forest concept is used to construct the NANTAR method for this proposed system. It is the mainstay of the recommended solution since it uses a machine learning model to classify network data and detect threats. The attacker nodes are subsequently isolated and RPL routing data is updated according to the classification results. The algorithm consists of the following steps: packet categorization, packet forwarding, and routing table updating. The NANTAR method uses the Random Forest model to categorize incoming packets as either normal or harmful based on their attributes. The NANTAR technique is used to extract features from the packet header, such as rank, hop count, parent, and destination, which are then put into the Random Forest model. In the section on Routing Table Update, the NANTAR method modifies the routing table in accordance with the RPL protocol if the packet is malicious; if it is not, it modifies the routing table to steer clear of the attacker node. The attacker node is eliminated from the candidate parent list and the neighbor list via the NANTAR algorithm, which also sets its rank and hop count to infinity. Using the most recent routing table as an example, the NANTAR algorithm forwards packets. The algorithm is as follows:

**Algorithm for NANTAR process**
**Require:** Network traffic data
**Ensure:** Harmful data are detected and ignored
**Initialization**
Variable declaration for data collection and preprocessing(v)
**In collection process:**
Declaring a threshold for a malicious node
**While** running the recall process **do**
Wait for timer event(tb== process-event-times)
**Received the timer event:**

Increment packet count( for counting the processed packet)
Access packet buffer to analyze the size of the packet
Accumulate packet size in total packet size
**Process the collected data**
Calculate the average-packet-size / packet-count
Perform machine learning model interference (average packet size):
**Return** if traffic is malicious
**If** malicious traffic **then**
Log-event("malicious traffic")
Block-traffic()
Adjust-network-traffic()
**End if**
Reset the buffer for the next step
**End while**
End the process

The perform Model Inference function is used by the algorithm to carry out machine learning model inference. If the traffic is found to be malicious, the function returns 1, and if it is benign, it returns 0. The algorithm calls the log Event function to log the event if malicious traffic is discovered. In this manner, the technique stops the attacker node from being selected as the parent or subsequent hop by the real nodes. Therefore, our goal is to evaluate the NANTAR algorithm's durability in the context of industry.

## 4. RESULT AND DISCUSSION

The result and discussion section will describe the reviews and analyze the NANTAR algorithm's performance. First, the accuracy score comparison of the models, which aids in identifying effective machine learning for this system

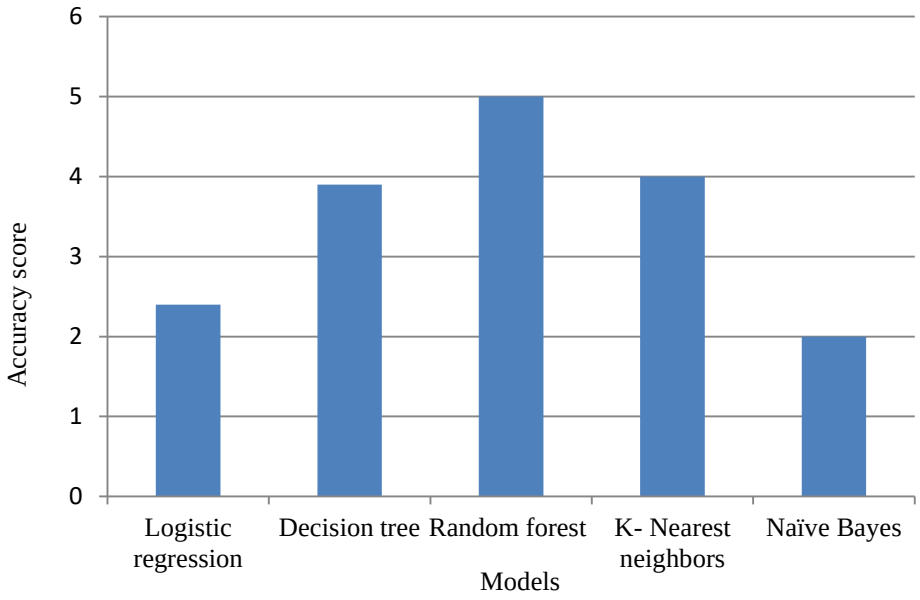


Figure 4. Evaluation for accuracy

The random forest approach offers a wider accuracy range than the other methods, as Figure 4 illustrates. Now assessing the NANTAR algorithm's performance. Objective Functions Zero (OF0) and Minimum Rank Hysteresis Objective Function (MRHOF), two historical Objective Functions (OFs), were compared with RPL with integrated NANTAR and regular RPL in order to undertake the evaluation. Metrics that are crucial for IoT security algorithms are used for the analysis set. Firstly throughput:

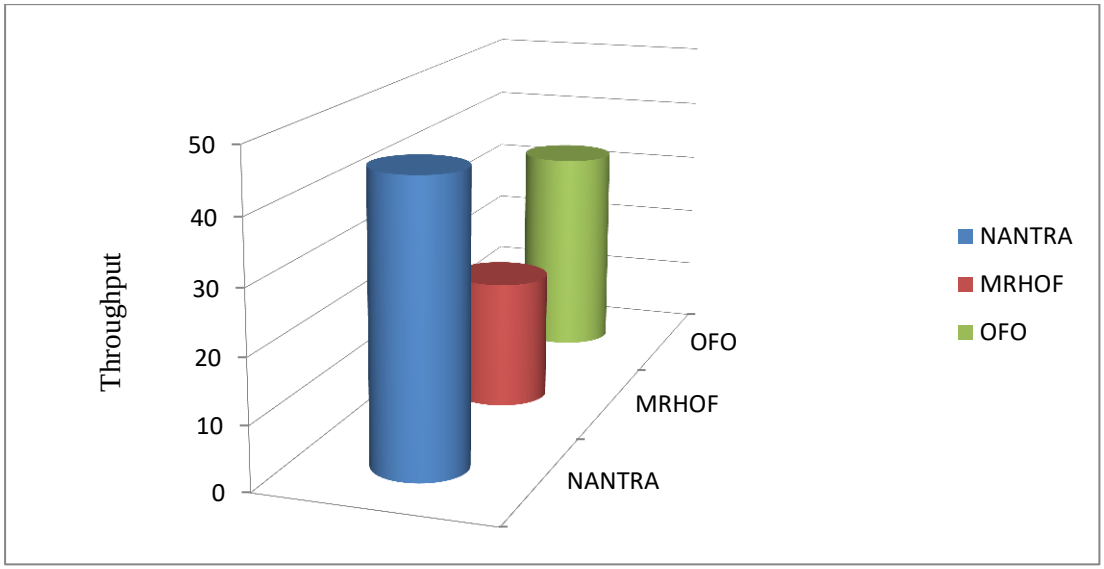


Figure 5. Comparing throughput.

According to Figure 5 , the NANTAR algorithm was able to process 45 packets per second, whereas MRHOF and OF0 were only able to process 22 and 26 packets per second, respectively. Optimized routing choices, real-time adaptation, anomaly detection, and machine learning effectiveness are the causes of this outcome. The models' latency is then compared.

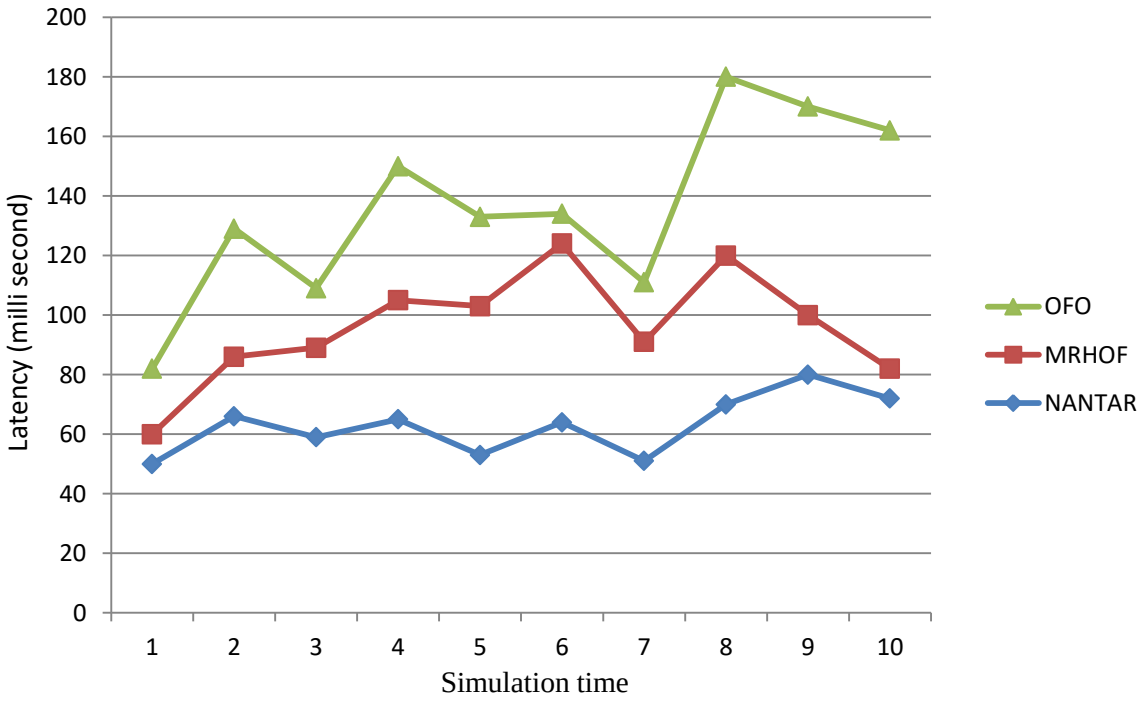


Figure 6. Latency Comparison

NANTAR achieved a 34% lower latency than MRHOF and a 26% lower latency than OF0, according to the data in Figure 6. The findings are supported by the fact that NANTAR's capabilities enable it to anticipate and steer clear of possible network bottlenecks. Additionally, the network is less likely to be slowed down by attack traffic when rogue nodes are promptly identified and isolated. This is because handling such threats requires additional processing, which can otherwise raise latency.

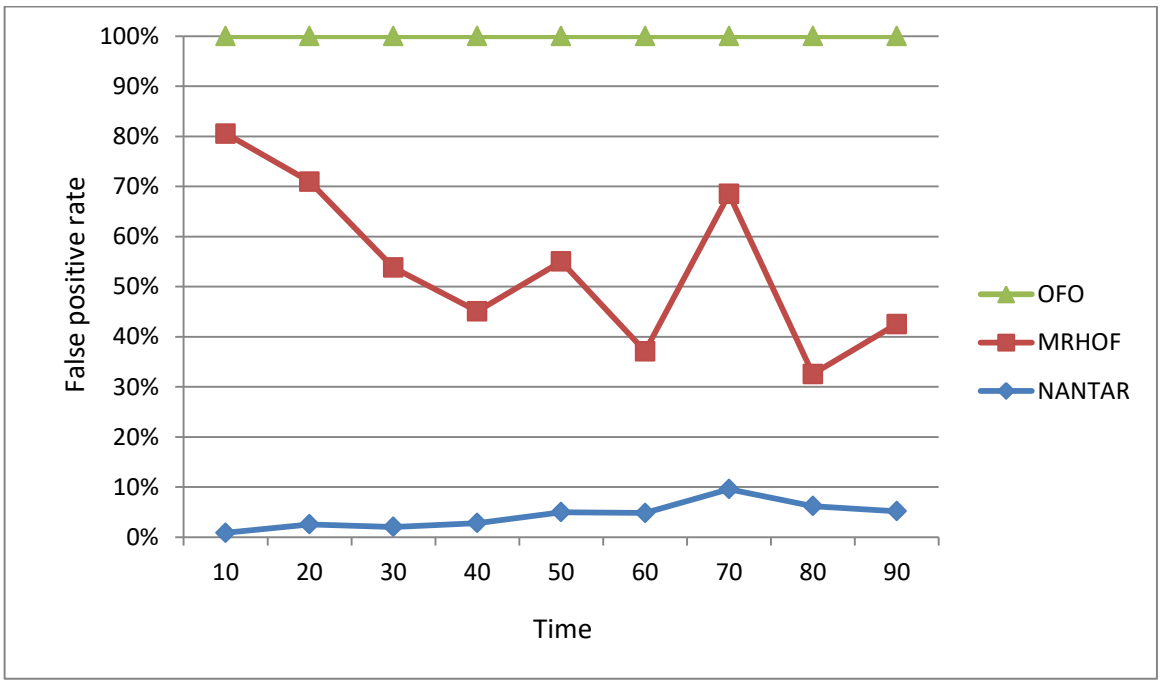


Figure 7. Comparing false positive rate

Figure 7 illustrates that NANTAR's performance remained constant as the network expanded. It also indicates that the efficacy and efficiency of the network stayed consistent even when the number of nodes increased.

The machine learning-based methodology of NANTAR allows it to detect intricate patterns and irregularities in network traffic. By using feature analysis that surpasses simple rule-based methods, NANTAR can distinguish between benign and malicious behavior more effectively. The NANTAR has the ability to dynamically adjust its detection thresholds based on the current circumstances. Its versatility allows it to balance false positives and false negatives to enhance overall security. Because of its ability to respond to detected anomalies, including isolating malicious nodes, NANTAR has lower error rates. Being quick reduces the likelihood of false negatives and prevents attacks from going unnoticed.

Lower False Positive Rate (FPR) and False Negative Rate (FNR) work together to help NANTAR maintain a balance between sensitivity and specificity. NANTAR is a suitable choice for IoT situations since it reduces both types of errors, improving overall network security.

## 5. CONCLUSION

This study introduced the NANTAR algorithm, which aims to enhance IoT routing within the framework of the Routing Protocol for Low-Power and Lossy Networks (RPL). By using AI techniques, NANTAR addresses significant problems including security vulnerabilities and subpar performance. These findings demonstrate considerable improvements in throughput, resource use, and latency. These results demonstrate significant gains in resource utilization, latency reduction, and performance. Furthermore, NANTAR's competitive power consumption and threat resistance make it a suitable option for safeguarding IoT environments. While promising, further research is clearly needed to determine its scalability. NANTAR's latency was 26% lower than MRHOF's and 34% lower than OF0's. NANTAR exhibited a 39% higher Throughput and an 18% lower false positive rate in comparison to MRHOF and OF0.

## REFERENCE

[1] Castro, O. E. L., Deng, X., & Park, J. H. ArticlesComprehensive Survey on AI-Based Technologies for Enhancing IoT Privacy and Security: Trends, Challenges, and Solutions.

[2] Raoof, A., Matrawy, A., & Lung, C. H. (2018). Routing attacks and mitigation methods for RPL-based Internet of Things. *IEEE Communications Surveys & Tutorials*, *21*(2), 1582-1606.

[3] Jhanjhi, N. Z., Brohi, S. N., Malik, N. A., & Humayun, M. (2020, October). Proposing a hybrid rpl protocol for rank and wormhole attack mitigation using machine learning. In *2020 2nd International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-6). IEEE.

[4] Alfriehat, N., Anbar, M., Aladaileh, M., Hasbullah, I., Shurbaji, T. A., Karuppayah, S., & Almomani, A. (2024). RPL-based attack detection approaches in IoT networks: review and taxonomy. *Artificial Intelligence Review*, *57*(9), 248.

[5] Verma, A., & Ranga, V. (2019, April). ELNIDS: Ensemble learning based network intrusion detection system for RPL based Internet of Things. In *2019 4th International conference on Internet of Things: Smart innovation and usages (IoT-SIU)* (pp. 1-6). IEEE.

[6] Jhanjhi, N. Z., Brohi, S. N., & Malik, N. A. (2019, December). Proposing a rank and wormhole attack detection framework using machine learning. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)* (pp. 1-9). IEEE.

[7] Cakir, S., Toklu, S., & Yalcin, N. (2020). RPL attack detection and prevention in the Internet of Things networks using a GRU based deep learning. *IEEE Access*, *8*, 183678-183689.

[8] Nayak, S., Ahmed, N., & Misra, S. (2021). Deep learning-based reliable routing attack detection mechanism for industrial Internet of Things. *Ad Hoc Networks*, *123*, 102661.

[9] Santos, C. L. D., Mezher, A. M., León, J. P. A., Barrera, J. C., Guerra, E. C., & Meng, J. (2023). Ml-rpl: Machine learning-based routing protocol for wireless smart grid networks. *IEEE Access*, *11*, 57401-57414.

[10] Niu, X. (2021). [Retracted] Optimizing DODAG Build with RPL Protocol. *Mathematical Problems in Engineering*, *2021*(1), 5579564.

[11] Zahra, F., Jhanjhi, N. Z., Brohi, S. N., Khan, N. A., Masud, M., & AlZain, M. A. (2022). Rank and wormhole attack detection model for RPL-based internet of things using machine learning. *Sensors*, *22*(18), 6765.

[12] Neerugatti, V., & Mohan Reddy, A. R. (2019). Machine learning based technique for detection of rank attack in RPL based internet of things networks. *Machine Learning Based Technique for Detection of Rank Attack in RPL based Internet of Things Networks (July 10, 2019). International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN*, 2278-3075.

[13] Al-Amiedy, T. A., Anbar, M., Belaton, B., Kabla, A. H. H., Hasbullah, I. H., & Alashhab, Z. R. (2022). A systematic literature review on machine and deep learning approaches for detecting attacks in RPL-based 6LoWPAN of internet of things. *Sensors*, *22*(9), 3400.

[14] Seyfollahi, A., & Ghaffari, A. (2021). A review of intrusion detection systems in RPL routing protocol based on machine learning for internet of things applications. *Wireless Communications and Mobile Computing*, *2021*(1), 8414503.

[15] Attique, D., Wang, H., & Wang, P. (2022). Fog-assisted deep-learning-empowered intrusion detection system for RPL-based resource-constrained smart industries. *Sensors*, *22*(23), 9416.