International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

Optimization Accuracy of Blackhole Attacks in Routing Protocol using Machine Learning

Akilesh Pavithran[1], Mr. Manish Sahu[2]

M. Tech. Scholar, Department of Electronics and Communication Engineering, SORT, People's University, Bhopal, India[1]

Assistant Professor, Department of Electronics and Communication Engineering, SORT, People's University, Bhopal, India[2]

Abstract

Blackhole attacks pose a significant threat to the reliability and stability of wireless ad hoc and sensor networks, as malicious nodes intentionally drop data packets after falsely advertising optimal routes. Such attacks lead to degradation in network performance by reducing the packet delivery ratio, increasing end-to-end delay, and causing severe routing disruptions. Traditional defense techniques rely on threshold-based or heuristic mechanisms, which often fail to adapt to dynamic network conditions and result in low detection accuracy. To address these limitations, machine learning (ML)-based approaches have emerged as an effective solution for detecting and mitigating Blackhole attacks. This paper proposes an optimized ML framework aimed at enhancing the accuracy of identifying Blackhole nodes within routing protocols such as AODV, DSR, and DSDV. The model analyzes essential network behavior features including abnormal sequence numbers, packet drop rate, routing overhead, and transmission patterns to classify nodes as benign or malicious. Various ML classifiers—such as Random Forest, SVM, Logistic Regression, KNN, and ensemble learning—are evaluated using performance metrics like accuracy, precision, recall, F1-score, and ROC curves.

Keywords: Routing Protocol, Blackhole Attack, Machine Learning

1. INTRODUCTION

Wireless ad hoc and sensor networks have gained significant importance due to their flexibility, scalability, and ability to operate without fixed infrastructure. These networks play a vital role in military communication, disaster recovery, environmental monitoring, smart cities, and IoT-based applications. However, the decentralized nature, dynamic topology, and open communication medium make them highly vulnerable to various security threats. Among these, the Blackhole attack is considered one of the most destructive routing-layer attacks. In a Blackhole attack, a malicious node deliberately advertises itself as having the shortest and most reliable path to the destination, thereby attracting data traffic. Once the packets are routed through this compromised node, it drops them intentionally, resulting in severe network disruption, reduced throughput, and degraded Quality of Service (QoS) [1, 2]. Traditional routing protocols such as AODV (Ad hoc On-Demand Distance Vector), DSR (Dynamic Source Routing), and DSDV (Destination-Sequenced Distance Vector) are highly susceptible to Blackhole attacks because they rely on trust-based route discovery mechanisms. These protocols do not inherently include authentication or anomaly detection features, making it easier for malicious nodes to exploit routing vulnerabilities. Conventional defense techniques—including watchdog mechanisms, threshold-based detection, and

cryptography—often suffer from limitations such as high computational overhead, inability to adapt to changing network conditions, increased energy consumption, and poor detection accuracy [3, 4].

With the rapid evolution of intelligent security solutions, Machine Learning (ML) has emerged as a promising approach to detect and mitigate Blackhole attacks effectively. ML models have the capability to learn hidden patterns in network behavior, enabling them to distinguish between normal and malicious nodes based on features such as abnormal sequence numbers, sudden route changes, high packet drop ratio, excessive route replies, and unusual transmission behavior. ML-based classifiers—including Support Vector Machines (SVM), Random Forest, Naïve Bayes, K-Nearest Neighbors (KNN), and ensemble learning algorithms—have shown superior performance in identifying malicious activity compared to conventional techniques [5].

To further enhance detection accuracy, optimization techniques such as Grid Search, Genetic Algorithms, Particle Swarm Optimization (PSO), and Bayesian Optimization are applied to fine-tune the hyperparameters of ML models. These optimization strategies help reduce false positives, improve predictive reliability, and increase the overall robustness of the detection system. Optimized ML models provide faster decision-making, reduced computational cost, and improved adaptability to dynamic network environments.

Given the increasing sophistication of security threats, there is a critical need for intelligent and optimized solutions to ensure secure communication in wireless networks. This research focuses on improving the accuracy of detecting Blackhole attacks by integrating machine learning with optimization techniques, ultimately contributing to more secure, efficient, and resilient routing protocols [6, 7].

2. RP

The RPL routing protocol for low power and lossy networks is a routing protocol designed specifically for wireless sensor networks [5, 6]. This protocol, especially, is crucial for applications such as the Internet of Things and Smart Grid. The objective of the routing protocol is to ensure energy-efficient, low-latency, and reliable communication in wireless sensor networks [7]. One of the main objectives of the RPL protocol is to route data packets between nodes with low power consumption, which extends the battery life of the nodes in the network and uses energy resources more efficiently [8]. RPL creates and manages network topologies at the node level to enable wireless sensor networks to transmit data efficiently. The protocol can create and manage various topology structures, such as using Destination Oriented Directed Acyclic Graphs (DODAG), a node-level hierarchical structure that ensures efficient data forwarding. The routing protocol aims to transmit data packets with low latency and high reliability while minimizing the energy consumption of nodes in the network.

2.1 ATTACKS IN RPL

RPL is vulnerable to a number of attacks such as external rank attack, internal rank attack (increased and decreased), version attack, Hello and DIS flooding attack, blackhole attack and selective forwarding attack. Among these attacks, some of the attacks resulted from modifying control packet fields like rank, version number, repair procedure and so on, which have a significant influence on the DODAG construction of RPL. All these attacks have been grouped under DAG inherited attacks. Remaining attacks are due to malicious nodes that impersonate as normal nodes and exhaust resources. These attacks have been grouped under Address inherited attacks.

Selective Forwarding attack: In this attack, the attackers after joining the DODAG forwards only the selective packets (including data packets) and drops the remaining, which degrade the performance of the network.
Blackhole attack: The attackers reject all incoming traffic (RPL control packets and data packets), causing the routing topologies to become unstable and limits the data packet transmissions to the root [9]. This attack has a more negative impact on non-storing mode of RPL.
Hello flooding attack: DIO packets in RPL are handled in the same way as HELLO packets. The attackers flood DIO packets which increases control overhead and hence depletes the energy of legitimate node [10].
Sinkhole attack: In this type of attack, the attackers advertise a fake route with superior metrics in an effort to make adjacent nodes to choose it as their preferred parent [11].

3. PROPOSED METHODOLOGY

An extremely basic method for making a far superior classifier is to total the expectations of every classifier and anticipate the class that gets the most votes. A hard voting classifier is the name given to this majority-vote classifier.
The following flow chart of proposed methodology is representing in fig. 1. Somewhat surprisingly, this voting classifier often achieves a higher accuracy than the best classifier in the ensemble. In fact, even if each classifier is a weak learner (meaning it does only slightly better than random guessing), the ensemble can still be a strong learner (achieving high accuracy), provided there are a sufficient number of weak learners and they are sufficiently diverse.
The **Random Forest Classifier** is a powerful ensemble machine learning algorithm widely used for classification tasks, including anomaly detection in networks. It operates by building multiple decision trees during training and combining their outputs for more accurate and robust predictions.

1. **High Accuracy**: Random Forest can handle complex relationships in data and works well with high-dimensional datasets.
2. **Feature Importance**: It ranks feature importance, helping to identify the most relevant factors for blackhole attack detection.
3. **Robustness**: It is less prone to overfitting compared to single decision trees.
4. **Scalability**: Suitable for large datasets, making it effective for IoT network environments with diverse traffic patterns.

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

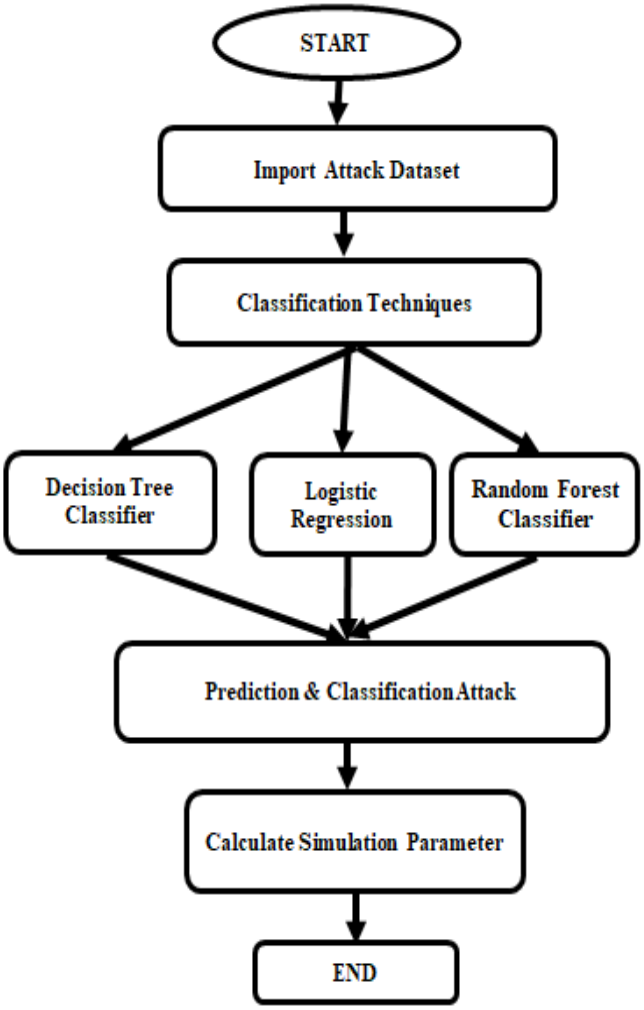

Figure 1: Flow chart of Proposed Methodology

The classifiers that we have utilized are LR, DT and RF.

Algorithm steps:

Input: D= {($\boldsymbol{x1}$,), ($\boldsymbol{x2}$,) ,…, ($\boldsymbol{xN}$,$\boldsymbol{yN}$)}, $\boldsymbol{L(y,O(x))}$

Where: ($\boldsymbol{y}$,($\boldsymbol{x}$)) is the approximate loss function.

Begin

Initialize: $(\boldsymbol{x}) = \frac{\boldsymbol{argmin}}{\boldsymbol{w}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{L(yi,w)}$

$\boldsymbol{for\ m}$=$\boldsymbol{1}$:$\boldsymbol{M}$

$$\boldsymbol{r_{im}} = -\frac{\boldsymbol{\partial L(yi,O(xi))}}{\boldsymbol{\partial O(xi)}}$$

Train weak learner $\boldsymbol{C_m(x)}$ on training data

Calculate $\boldsymbol{w}$: $\boldsymbol{w_m} = \boldsymbol{\arg\min} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{L(yi, O_{m-1}(xi) + wC_m(xi))}$

Update : $\boldsymbol{O_m(x)} = \boldsymbol{O_{m-1}(x) + wC_m(x)}$

End for

End

Output: $\boldsymbol{O_m(x)}$

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

4. SIMULATION RESULTS

Simulation Parameter

Accuracy gives a proportion of how precise your model is in anticipating the real up-sides out of the absolute up-sides anticipated by your framework. Review gives the quantity of real up-sides caught by our model by grouping these as obvious positive. F-measure can give a harmony among accuracy and review, and it is liked over precision where information is uneven.

Accordingly, F-measure was used in this review as a presentation metric to give a decent and fair measure utilizing the equation.

$$Accurancy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\Pr ecision = \frac{TP}{TP + FP} \quad (2)$$

$$\mathrm{Re}\, call = \frac{TP}{TP + FN} \quad (3)$$

$$F - Score = \frac{2(\Pr ecision \times \mathrm{Re}\, call)}{\Pr ecision + \mathrm{Re}\, call} \quad (4)$$

Where,
TP = True Positive,
TN = True Negative
FP = False Positive,
FN = False Negative

The provided data seems to represent the distribution of different attack types in a dataset. Here's a description of the data:

Attack type	
Normal	340066
Grayhole	14596
Blackhole	10049
TDMA	6638
Flooding	3312

Figure 2: Attack Types

- **Normal**: This category has the highest count, with 340,066 occurrences. It likely refers to the baseline or normal traffic that is not associated with any attacks.

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

- **Grayhole**: This type of attack is represented by 14,596 occurrences. A grayhole attack typically involves selectively dropping packets, causing communication disruptions while pretending to be functioning properly.
- **Blackhole**: There are 10,049 instances of this attack. In a blackhole attack, malicious nodes drop all incoming packets, causing a denial of service without forwarding any of the packets.
- **TDMA (Time Division Multiple Access)**: This category has 6,638 occurrences. TDMA attacks could involve exploiting the time-slot allocation in communication systems to disrupt or interfere with transmissions.
- **Flooding**: There are 3,312 instances of a flooding attack. Flooding typically refers to overwhelming the network or system with excessive traffic, often seen in denial-of-service (DoS) attacks.

Pre-processing –

1. Checking Null values df. isnull().sum()
2. **Standard Scaler**: This technique standardizes features by removing the mean and scaling to unit variance. It ensures that each feature contributes equally to model training, especially for algorithms that are sensitive to feature scaling (e.g., logistic regression, SVM).
3. **Label Encoder**: Label Encoder encodes categorical labels into numerical values, converting each unique category into an integer. It is useful for transforming target labels in classification tasks.
4. **OneHot Encoder**: OneHot Encoder converts categorical features into a series of binary columns, where each column represents one possible category. This is ideal for handling non-ordinal categorical variables in machine learning models.

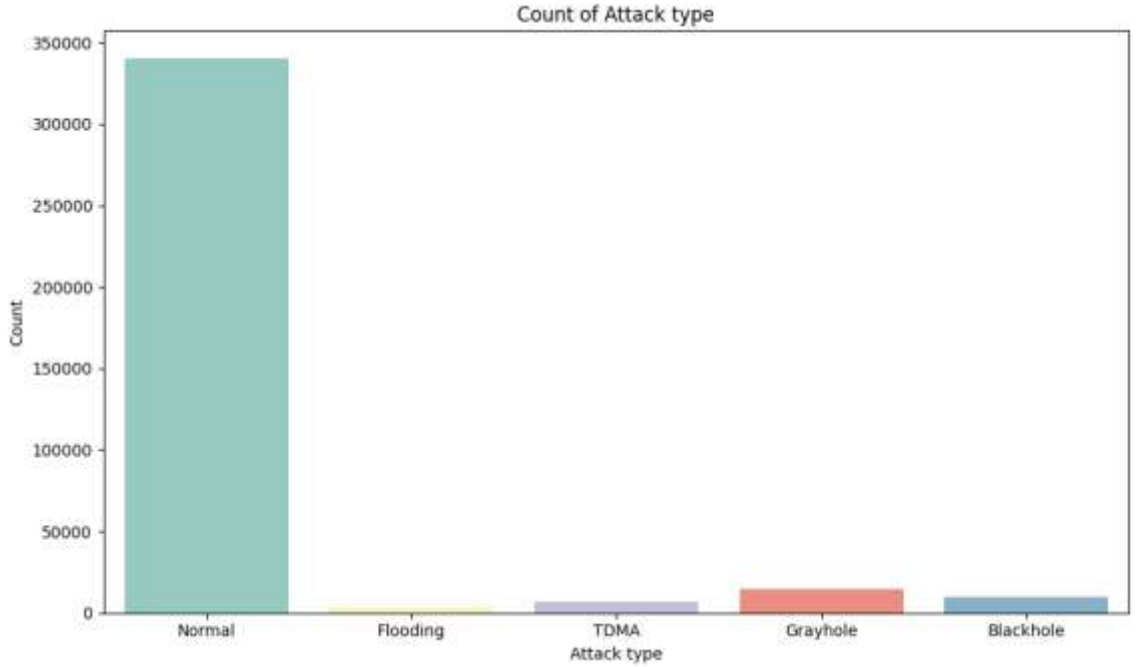

Figure 3: Graph of Attack Type

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

The table presents the performance metrics (Accuracy, Precision, Recall, and F1-score) for three machine learning models: Logistic Regression, Decision Tree, and Random Forest. The **Logistic Regression** model achieved an accuracy of 0.9594 with relatively balanced precision and recall, leading to an F1-score of 0.9621. The **Decision Tree** model demonstrated high performance, with an accuracy, precision, recall, and F1-score of 0.9935. The **Random Forest** model outperformed the others, achieving the highest scores across all metrics with an accuracy, precision, recall, and F1-score of 0.9962. Overall, Random Forest leads in all aspects, followed by Decision Tree and Logistic Regression.

Table 1: Comparison Result

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.9594	0.9674	0.9594	0.9621
Decision Tree	0.9935	0.9935	0.9935	0.9935
Random Forest	0.9962	0.9962	0.9962	0.9962

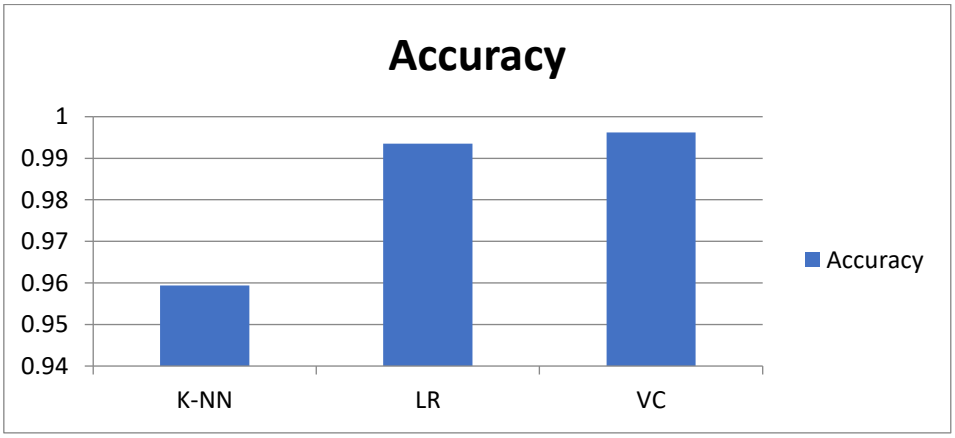

Figure 4: Graphical Represent of Accuracy

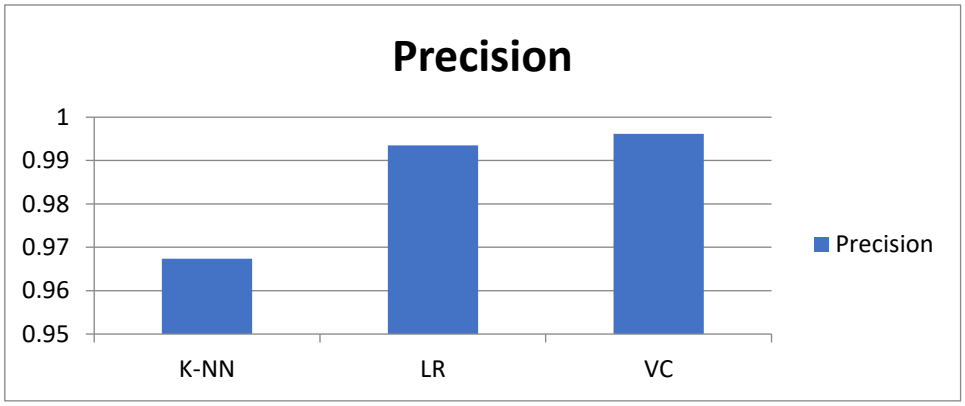

Figure 5: Graphical Represent of Precision

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

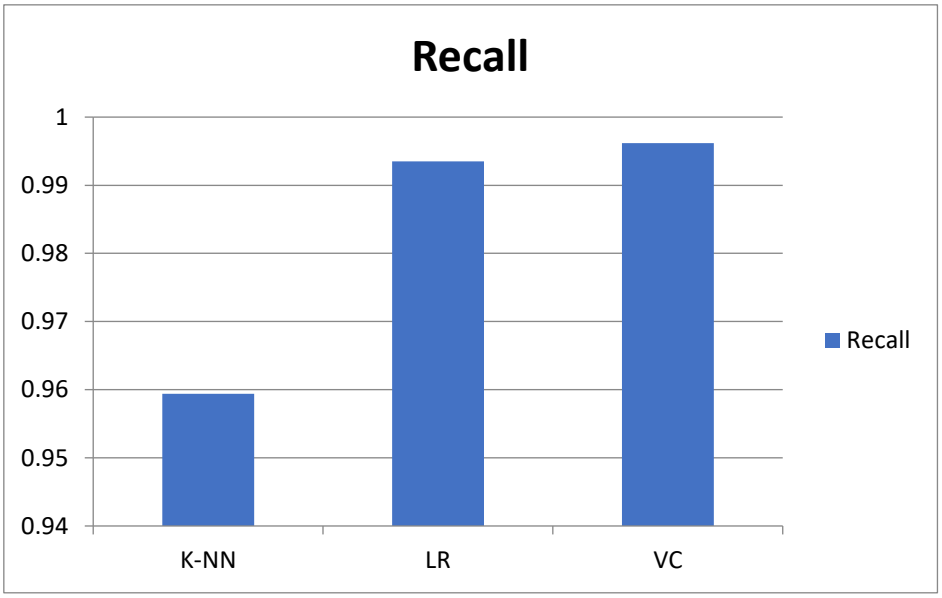

Figure 6: Graphical Represent of Recall

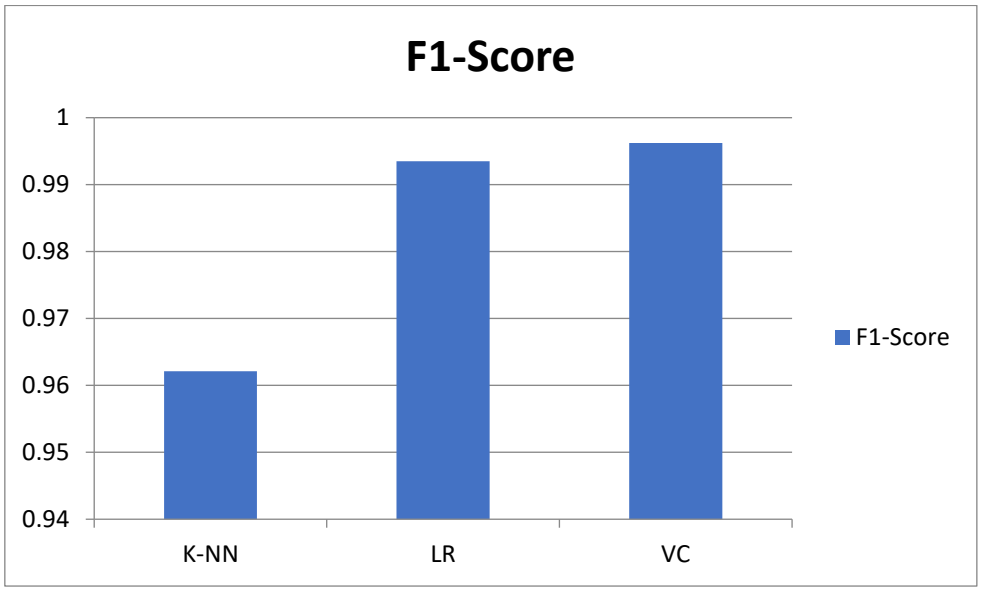

Figure 7: Graphical Represent of F1-Score

5. CONCLUSIONS

Blackhole attacks remain one of the most critical challenges in wireless ad hoc and sensor networks, as they significantly disrupt routing operations and degrade overall network reliability. Traditional security mechanisms are often inadequate in detecting such attacks due to their static nature, limited adaptability, and high computational overhead. In this study, machine learning-based detection techniques are explored as an effective approach to enhance the identification and mitigation of Blackhole attacks in routing protocols such as AODV, DSR, and DSDV. By analyzing key network features—such as packet drop ratio, abnormal sequence numbers, routing overhead, and sudden route request variations—ML models can efficiently differentiate between normal and malicious nodes.

This research underscores the potential of machine learning as a scalable and intelligent defense mechanism against Blackhole attacks. By optimizing model parameters and leveraging advanced classification techniques, network security can be significantly strengthened. The findings serve as a foundation for developing real-time intrusion detection systems capable of protecting next-generation wireless networks from increasingly sophisticated routing-layer threats. Future work may involve integrating deep learning architectures, real-time detection modules, and hybrid security frameworks to further enhance the resilience of wireless communication systems.

REFERENCES

[1] F. A. Garba, O. Oduwole, A. B. Isa, A. S. Aliyu and R. M. Dima, "Comprehensive Analysis of Blackhole Attack on RPL in Static, Mobile and Hybrid Network Environment," *2024 IEEE 5th International Conference on Electro-Computing Technologies for Humanity (NIGERCON)*, Ado Ekiti, Nigeria, 2024, pp. 1-5.

[2] H. B H, A. Venkata Mandalam and A. Giri, "HVSNA: An Advanced Hybrid Attack on RPL-Based Low-Power Wireless Networks," *2023 IEEE North Karnataka Subsection Flagship International Conference (NKCon),* Belagavi, India, 2023, pp. 1-6

[3] Philokypros P. Ioulianou, Vassilios G. Vassilakis, Siamak F. Shahandashti, "ML-based Detection of Rank and Blackhole Attacks in RPL Networks", 13th International Symposium on Communication Systems, Networks and Digital Signal Processing, IEEE 2022.

[4] B. Aydın, S. Görmüş, H. Aydın, and S. Kulcu, "A new routing objective function for ietf 6tisch protocol," in 2022 30th Signal Processing and Communications Applications Conference (SIU). IEEE, 2022, pp. 1–4.

[5] P. P. Ioulianou, V. G. Vassilakis, and S. F. Shahandashti, "A trust-based intrusion detection system for RPL networks: Detecting a combination of rank and blackhole attacks," Journal of Cybersecurity and Privacy, vol. 2, no. 1, pp. 124–153, March 2022.

[6] A. Agiollo, M. Conti, P. Kaliyar, T.-N. Lin, and L. Pajola, "Detonar: Detection of routing attacks in rplbased iot," IEEE transactions on network and service management, vol. 18, no. 2, pp. 1178–1190, 2021.

[7] B. Farzaneh, M. Koosha, E. Boochanpour, and E. Alizadeh, "A new method for intrusion detection on rpl routing protocol using fuzzy logic," in 2020 6th International Conference on Web Research (ICWR), 2020, pp. 245–250.

[8] E. G. Ribera, B. M. Alvarez, C. Samuel, P. P. Ioulianou, and V. G. Vassilakis, "Heartbeat-based detection of blackhole and greyhole attacks in RPL networks," in 12th IEEE/IET International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Porto, Portugal, July 2020, pp. 1–6.

[9] R. Smith, D. Palin, P. P. Ioulianou, V. G. Vassilakis, and S. F. Shahandashti, "Battery draining attacks against edge computing nodes in IoT networks," Cyber-Physical Systems, vol. 6, no. 2, pp. 96–116, January 2020.

[10] C. Samuel, B. M. Alvarez, E. G. Ribera, P. P. Ioulianou, and V. G. Vassilakis, "Performance evaluation of a wormhole detection method using round-trip times and hop counts in RPL-based 6LoWPAN networks," in 12th IEEE/IET International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Porto, Portugal, July 2020, pp. 1–6.

[11] U. Singh, M. Shukla, A. K. Jain, M. Patsariya, R. Itare and S. Yadav, *Trust-Based Model for Mobile Ad-Hoc Network in the Internet of Things*, vol. 98, 2020.

[12] P. P. Ioulianou and V. G. Vassilakis, "Denial-of-service attacks and countermeasures in the RPL-based Internet of Things," 2nd International Workshop on Attacks and Defenses for Internet-of-Things (ADIoT) in conjunction with ESORICS, Luxemburg, Sept. 2019.

[13] P. P. Ioulianou, V. G. Vassilakis, and M. D. Logothetis, "Battery drain denial-of-service attacks and defenses in the Internet of things," Journal of Telecommunications and Information Technology, vol. 2, pp. 37–45, April 2019.

[14] A. Verma and V. Ranga, "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). IEEE, 2019, pp. 1–6.

[15] E. Aydogan, S. Yilmaz, S. Sen, I. Butun, S. Forsström, and M. Gidlund, "A central intrusion detection system for rpl-based industrial internet of things," in 2019 15th IEEE International Workshop on Factory Communication Systems (WFCS). IEEE, 2019, pp. 1–5.