Journal of Computer Science

Original Research Paper

Enhancing IoT RPL Protocol Security Against Black Hole Attacks with Deep Learning Techniques

[1]Krari Ayoub, [1]Hajami Abdelmajid, [1]Toubi Ayoub and [2]Mihi Soukaina

[1]Laboratory of Research Watch for Emerging Technologies (VETE), Hassan First University of Settat, Morocco
[2]Laboratory of Computer, Networks, Mobility and *Modeling (IR2M), Hassan First University of Settat, Morocco*

Article history
Received: 07-03-2024
Revised: 31-05-2024
Accepted: 22-06-2024

Corresponding Author:
Krari Ayoub
Laboratory of Research Watch for Emerging Technologies (VETE), Hassan First University of Settat, Morocco
Email: ayoub.krari@uhp.ac.ma

Abstract: As the Internet of Things (IoT) extends its reach into critical infrastructures, it encounters escalating security threats, particularly black hole attacks that jeopardize network communication. The growing use of IoT in essential services underscores the need for robust network integrity. Driven by this urgency, our research has developed a sophisticated detection framework that effectively identifies black hole attacks within IoT networks utilizing the Routing Protocol for Low-Power and Lossy Networks (RPL). Our approach leverages the advanced pattern recognition capabilities of a deep Multi-Layer Perceptron (MLP) model, which has been rigorously trained and validated on a dataset generated under simulated conditions in Cooja. Key accomplishments of our study include achieving a high overall model accuracy of 94.3%, with specific accuracies of 94.4% for training, 94.2% for validation, and 94.0% for testing. The model exhibited minimal Mean Squared Error (MSE) values, with the lowest recorded validation MSE at approximately 0.029192. Additionally, the model's performance was marked by nearly perfect Receiver Operating Characteristic (ROC) curves, demonstrating Areas Under the Curve (AUC) close to 1 for both classes across all datasets. These performance metrics validate the model's efficacy in discerning the subtleties of black hole attacks, thereby enhancing network security analytics and contributing significantly to the proactive defense mechanisms against cyber threats in IoT networks. Our findings not only demonstrate the capabilities of deep learning models in cybersecurity but also underscore the importance of innovative solutions in safeguarding the expanding landscape of IoT infrastructures.

Keywords: IoT-Enabled Devices, IoT Security, RPL Protocol, Black-Hole Attack, Deep Learning

Introduction

The exponential growth of the Internet of Things has ushered in an era of unparalleled connectivity. This integration of diverse devices has led to significant advancements in data collection and analysis, resulting in improved decision-making and enhanced efficiency across various sectors (Mamdiwar *et al*., 2021). However, this expanded network has also introduced complex security challenges, notably black hole attacks, which pose a severe threat to the reliability and confidentiality of IoT networks (Reshi *et al*., 2024). These attacks result in the disruption of service and potential privacy breaches, Traditional security measures often fall short in detecting and preventing such sophisticated attacks, due to their static nature and inability to adapt to the dynamic landscape of IoT threats (Muzammal *et al*., 2021). Machine and deep learning techniques offer a dynamic and flexible approach, enabling the automatic identification of unusual network patterns that signify potential threats. These advanced methodologies can adapt over time to changing network conditions and evolving attack strategies, positioning them as superior alternatives to rule-based systems (Aldweesh *et al.*, 2020).

The present study introduces an innovative framework that significantly enriches the field of IoT security through the application of a deep Multi-Layer Perceptron (MLP) model. This research distinguishes itself by not only recognizing the patterns indicative of black hole attacks but also by adapting to the network's evolving conditions in real time. The proposed approach advances the current state-of-the-art by providing a more resilient and robust solution that enhances the detection accuracy of black hole attacks in IoT networks.

Materials and Methods

In confronting the intricate security challenges posed by black hole attacks within IoT networks, research in this domain has progressively evolved to address these threats effectively. Our meta-analysis scrutinizes pivotal contributions to this field, thoroughly detailing the methodologies, metrics, and outcomes employed by various studies, while also acknowledging the limitations of each. This comprehensive review not only encapsulates the progression of security strategies but also introduces our innovative application of deep learning to enhance attack detection within RPL-based IoT frameworks. Table (1) below succinctly captures the essence of these developments, providing a foundational context for the introduction of our proposed deep Multi-Layer Perceptron (MLP) model.

RPL Black-Hole Attack

An RPL black hole attack (Fig. 1) specifically targets the Routing Protocol for Low-Power and Lossy Networks (RPL), which is crucial for the operational functionality of Internet of Things (IoT) networks. RPL optimizes communication among nodes by establishing efficient paths for data transmission to a central sink node. During such an attack, an assailant impersonates a legitimate node, gaining trust and thereby compromising the network. This compromised position is exploited to selectively disrupt the network by dropping key RPL control messages, such as the DODAG Information Object (DIO), Destination Advertisement Object (DAO), and DODAG Information Solicitation (DIS) messages. These messages are essential for maintaining and updating the Directed Acyclic Graph (DAG) topology, which is vital for routing data packets efficiently.

Table 1: Advancements in RPL IoT network defense

Study	Year	Methodology	Key metrics	Outcome	Limitations
Neerugatti and Mohan Reddy (2019)	2019	MLTKNN based on K-nearest neighbor	True positive rate, false positive rate, delay, packet delivery rate	Improved security and performance in IoT networks	Limited to simulation environment, may not reflect real-world complexities; no deep learning techniques used; no black hole attack addressed; limited simulation data and limited analysis
Kamis *et al*. (2023)	2023	RPL, IPv6 over 6LoWPAN	Denial of Service (DoS) attacks,	Explored challenges and protection	No empirical data; theoretical analysis
Ezzitouni *et al*. (2021)	2021	SecRPL-MS with quantum inspired Neural network and prince algorithm	Malicious node detection accuracy, security metrics	Enhanced security in IoT networks against multiple attacks	Complex setup may limit practical deployment; addressed routing but not black hole attacks; energy-consuming node mechanism
Rakesh and Parveen Sultana (2023)	2023	E-RAD algorithm	Energy consumption, packet delivery delay, packet delivery accuracy	Reduced energy consumption and improved packet delivery metrics	Specific to rank attacks, may not address other vulnerabilities; no black hole attack addressed; no deep learning techniques used; node energy consumed using the proposed algorithm
Nandhini *et al*. (2023)	2023	Simulation in cooja for RPL	Packet overhead, inter-packet time, packet delivery ratio, power consumption	Highlighted adverse effects, and potential countermeasures against rank attacks	Focus on rank attacks may not address other vulnerabilities; no black hole attack addressed; no deep learning techniques used; node energy consumed using the proposed algorithm
Bang and Rao (2023)	2023	Deep Multi-Layer Perceptron (MLP) model	Model accuracy, loss ROC, confusion matrix	Accuracy of 94.3%, with specific accuracies of 94.4% for training, 94.2% for validation, and 94.0% for testing	-
Our Proposed Approach	2024				

The attacker's actions prevent the network from maintaining its structure and routing data effectively, causing significant communication failures among legitimate nodes. These nodes struggle to identify optimal data delivery paths to the sink node, resulting in substantial data loss and inefficient use of network resources. Additionally, the attacker may opt to selectively drop data packets or entirely discard all received packets, further amplifying the network disruption (Deepavathi and Mala, 2023).

The ramifications of an RPL black hole attack are severe, extending beyond simple communication disruption to substantial data loss, resource wastage, and notable degradation of network performance. Attackers might exploit such disruptions to extract sensitive data or initiate more severe, complex cyber-attacks (Arshad *et al*., 2022).

Detecting an RPL black hole attack is challenging, primarily because the attack involves omission, rather than alteration, of data packets. Conventional security mechanisms that detect corrupted or modified packets are inadequate. Detection strategies must instead scrutinize the network's control message traffic. Monitoring inconsistencies in the frequency, delivery, and integrity of DIO, DAO, and DIS messages can reveal disruptions in the DAG topology, indicating a potential black hole attack. This necessitates advanced monitoring tools capable of analyzing control message flow in real time and detecting patterns that suggest tampering (Alazab *et al*., 2023).

In conclusion, an RPL black hole attack constitutes a sophisticated Denial of Service (DoS) attack exploiting the vulnerabilities of the RPL protocol within IoT networks. Developing and implementing effective mitigation and detection strategies is crucial to preserve the integrity and functionality of IoT networks that depend on this protocol.

Proposed Approach

Our framework (Fig. 2) uses deep learning to detect malicious nodes in IoT networks. We simulate network traffic, preprocess and normalize the data, and then use PCA for feature extraction. An MLP with three hidden layers and the Adam optimizer classifies nodes, using k-fold cross-validation to ensure accuracy. This method effectively detects black-hole attacks, maintaining network integrity.

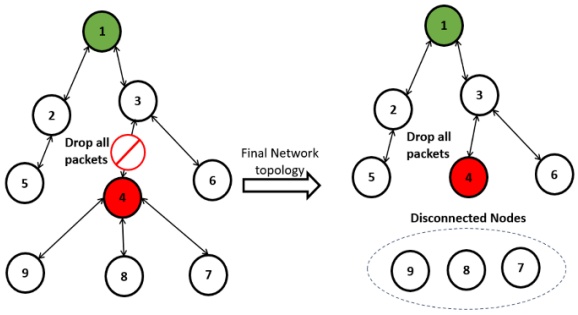

Fig. 1: Black hole attack process

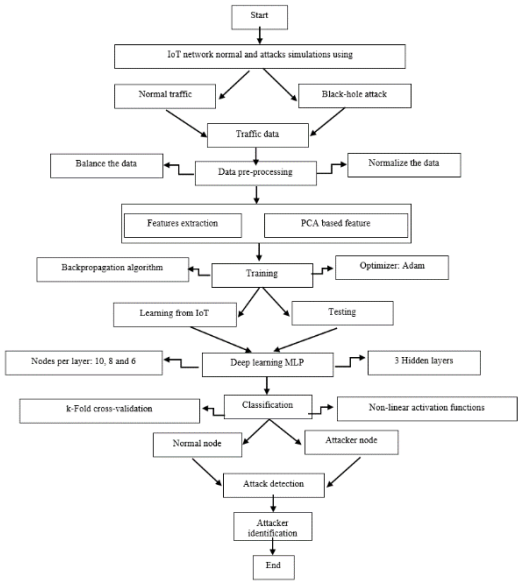

Fig. 2: Proposed approach

Normal and Attack Simulations

In the course of developing our innovative approach to detect black hole attacks in IoT networks, we undertook a meticulous process of data collection, transformation, and analysis. This process began with conducting simulations as shown in Figs. (3-4) and table configuration (Table 2) that mimic the conditions of a black hole attack to accumulate a dataset reflective of malicious network behavior. Parallel to this, simulations representing the normal operational behavior of nodes were executed to compile a benign dataset. These simulations were crucial for capturing the nuanced differences between compromised and uncompromised network states.

The data harvested from these simulations were initially captured in PCAP format using the Cooja simulator, an integral component of the Contiki OS designed specifically for IoT network simulations (Khan *et al*., 2023). These PCAP files were then transformed into a more analytically friendly CSV format using Wireshark, facilitating a seamless transition to the data preparation phase. Employing robust Python libraries, namely NumPy and pandas, the raw data underwent a comprehensive cleaning and preprocessing regimen. This preparation involved coding, labeling and ultimately partitioning the dataset into distinct sets for training and testing purposes.

The prepared dataset served as the input for our neural network-based model, specifically a deep Multi-Layer Perceptron (MLP) model, illustrated in Fig. (2). The MLP model stands at the core of our detection framework, leveraging its deep learning capabilities to discern patterns indicative of black hole attacks amidst complex network traffic data. This deep MLP model embodies the culmination of our preparatory work, poised to scrutinize network behavior for anomalies that signal malicious interference.

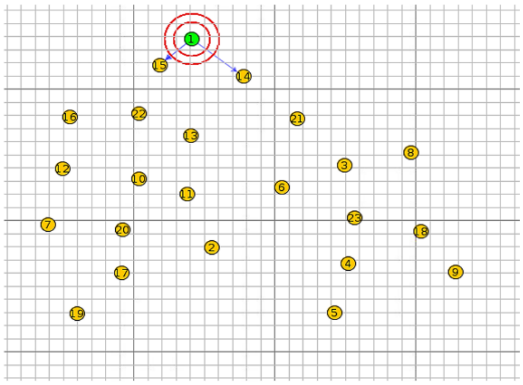

Fig. 3: Normal simulation map

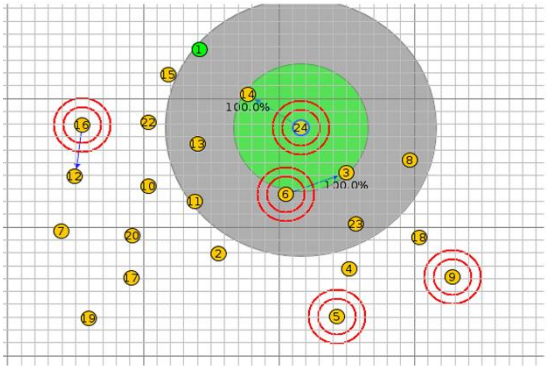

Fig. 4: Black hole attack simulation map

Table 2: Simulations configuration and parameters

Parameters	Values
Node type	SKY mote
Os version	Contiki 3.0
Protocol	RPL
Radio medium	Unit disk graph Medium: distance loss
Objective function	MRHOF
TX range	50/100 m
Interface range	50/100 m
Simulation area	100×100 m
MTU size	1280 Byte
Simulation duration	120 min
Sender nodes	23
Sink node	1
Repetitions	5

Sensor Maps

Upon close inspection of the sensor maps (Figs. 5-6) representing an IoT network, the ramifications of a black hole attack are readily apparent. The pre-attack state of the network showcases a well-structured mesh of sensor nodes, with bidirectional communication paths that signify an ideal operational network. Each node is denoted by a pair of numerical values, suggesting a measurement of network-specific metrics. The post-attack sensor map, however, reveals a network compromised by the strategic failure of critical nodes, as evidenced by the cluster of eight disconnected nodes. This disconnection is symptomatic of the black hole attack's potency, wherein the malicious entity succeeds in intercepting and nullifying the data packets destined for these nodes, effectively severing their communication links (Sasi *et al.*, 2023).

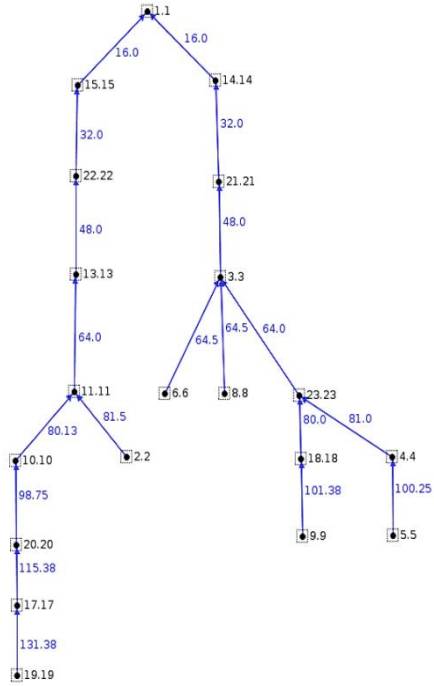

Fig. 5: Final sensor map of the normal simulation

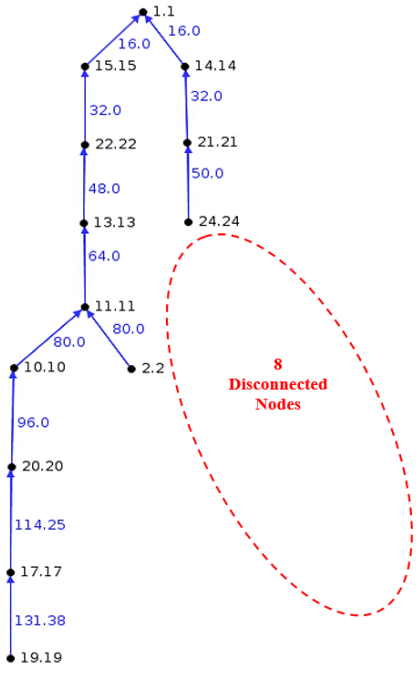

Fig. 6: Final sensor map of the attack simulation

DOI: 10.3844/jcssp.2024.1530.1544

The disconnected nodes represent a significant portion of the network, indicating a high attack severity. Their isolation from the network's continuum not only impedes the flow of data but also poses a strategic risk, potentially enabling the attacker to exert control over the flow of information or to leverage the compromised nodes for further attacks. The quantitative impact of this disconnection can be inferred from the substantial alteration in the numerical values associated with the affected nodes, which may reflect a degradation in signal quality, an increase in packet loss, or a depletion in resource availability. These metrics underscore the need for a vigilant security framework capable of early detection and response to such covert attacks.

The graphs in (Figs. 7-8) show the power consumption of nodes in the network during normal conditions and under a black hole attack, as simulated in the Cooja simulator. In the normal scenario, the power consumption fluctuates but remains mostly within a lower range. However, under the black hole attack, there are noticeable spikes in power consumption, indicating abnormal behavior and possibly the additional effort required by the network to deal with the malicious activity. These anomalies could be key indicators for the MLP deep learning model to detect and flag potential black hole attacks. The patterns observed in the power consumption during the black hole attack simulation are distinct from the normal power usage patterns, which could be utilized to train machine or deep learning models for effective anomaly detection (Krari *et al*., 2023).

The bar graphs in (Figs. 9-10) present instantaneous power consumption in a network under normal conditions and during the black hole attack. Each bar represents a node, with segments indicating power used for the Low-Power Mode (LPM), Central Processing Unit (CPU), radio listening, and radio transmission. Under normal conditions, the consumption is lower and relatively uniform across the nodes. In contrast, during a black hole attack, there is a notable increase in power used for radio transmission, suggesting an increase in network traffic as nodes attempt to route data through the compromised network. This change in power distribution serves as a valuable indicator for detecting anomalies associated with black hole attacks.

The provided bar graphs shown in Figs. (11-12) illustrate the average radio duty cycle for nodes in a network during typical operation and under a black hole attack. The duty cycle represents the percentage of time a node's radio is active, either listening or transmitting. In both scenarios, the duty cycle for radio transmission (blue bars) is notably higher during the black-hole attack, indicating increased network traffic as nodes respond to the malicious activity. This consistent increase across the nodes can be an effective metric for the deep learning Model to detect and address black hole attacks (Krari *et al*., 2021).

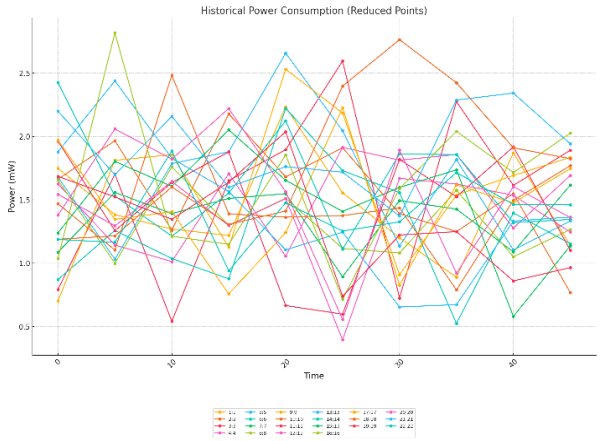

Fig. 7: Historical power consumption during normal simulation

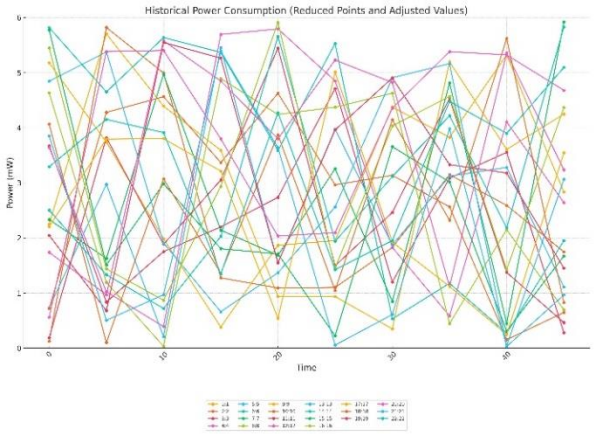

Fig. 8: Historical power consumption during black hole attack

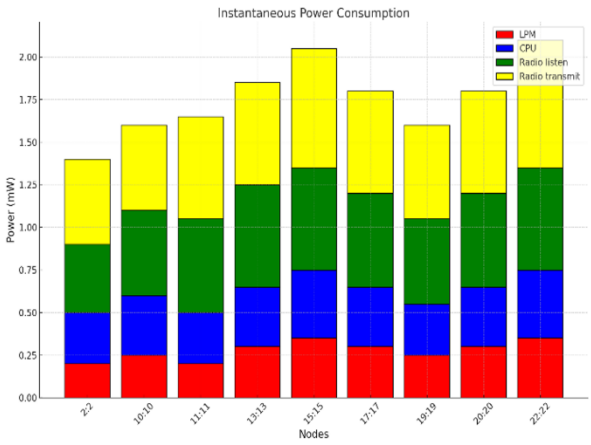

Fig. 9: Instantaneous power consumption during normal simulation

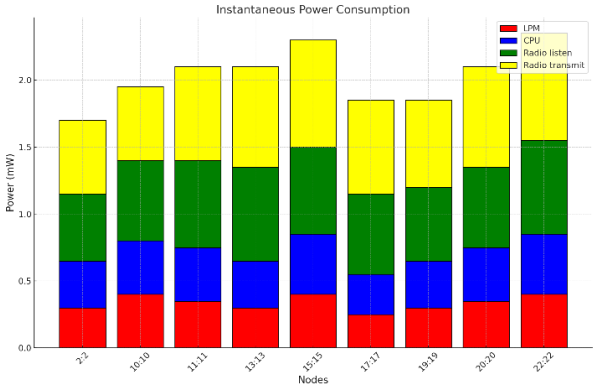

Fig. 10: Instantaneous power consumption during black hole attack

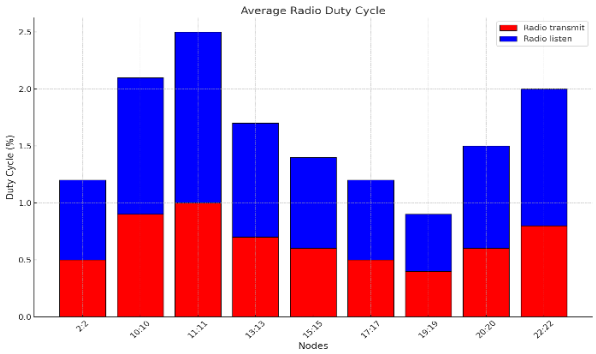

Fig. 11: Average radio duty cycle consumption during normal simulation

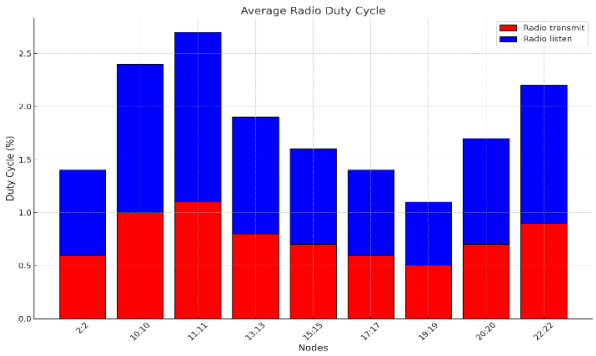

Fig. 12: Average radio duty cycle during black hole attack

During the normal simulation as shown in Fig. (13) and the simulated black hole attack, the graph Fig. (14) on the right exhibits both anomalies in neighbor connections and evidence of node disconnections. While certain nodes show an unexpected increase in neighbor counts, indicative of the attack's influence, others have no reported data, suggesting they were severed from the network. This loss of connectivity highlights the attack's severity, disrupting the network to an extent where data collection from some nodes became impossible, providing a stark contrast to the left graph's stable network conditions under normal operation.

Fig. 13: Average neighbor count during normal simulation

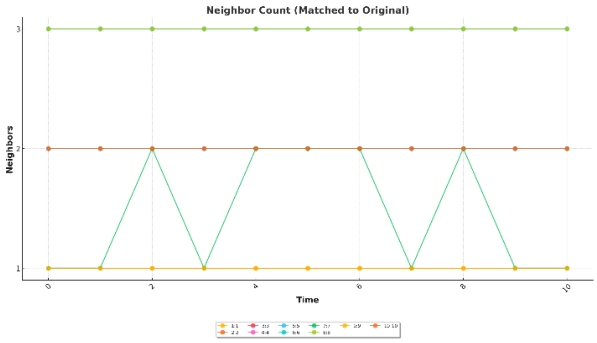

Fig. 14: Average neighbor count during black hole attack

The graphs in Figs. (15-16), corresponding to the normal and black hole attack simulation shows substantial disruption in beacon intervals across network nodes. The intervals are dramatically varied, with some intervals elongating significantly, directly indicating the network's compromised state. Such aberrations in beacon timings reflect the network's response to the malicious entity, as nodes struggle to maintain communication amidst the introduced chaos, unequivocally signaling the presence of a black hole attack (Arshad *et al.*, 2021).

The figures depicted in Figs. (17-18) provide a comprehensive visual representation of the repercussions observed in network performance stemming from both normal simulation conditions and the disruptive presence of a black hole attack. A discernible pattern emerges wherein there is a substantial escalation in the average routing metric fluctuation, coupled with a striking diminution in the successful transmission of packets.

Notably, the impact of the black hole attack becomes evident through a conspicuous decline in packet delivery rates. This deleterious effect can be attributed to the insidious nature of the attack, which involves the deliberate manipulation or outright dropping of packets by malicious entities. Consequently, this nefarious activity necessitates an increased frequency of routing attempts and introduces heightened intervals in the network operation.

In essence, the graphical representation serves as a quantitative illustration of the erosion of network integrity caused by the pernicious black hole attack. It underscores the imperative for robust security measures to safeguard against such threats and underscores the critical importance of fortifying network defenses to preserve operational efficacy and reliability.

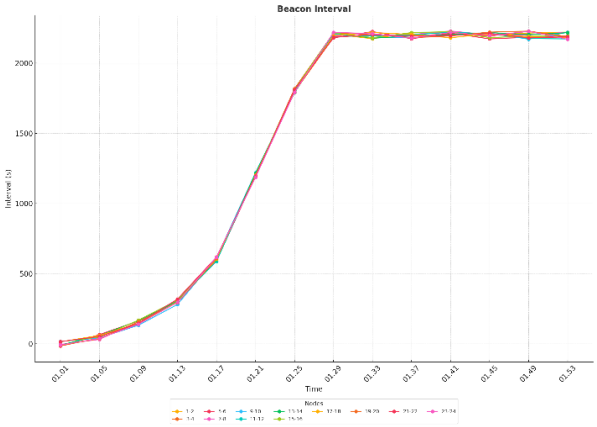

Fig. 15: Beacon interval during normal simulation

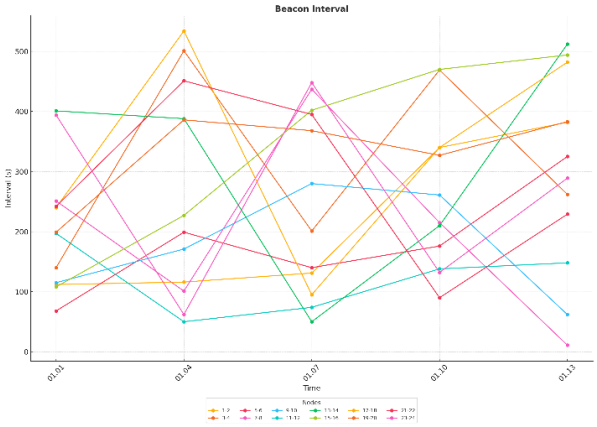

Fig. 16: Beacon interval during black hole attack

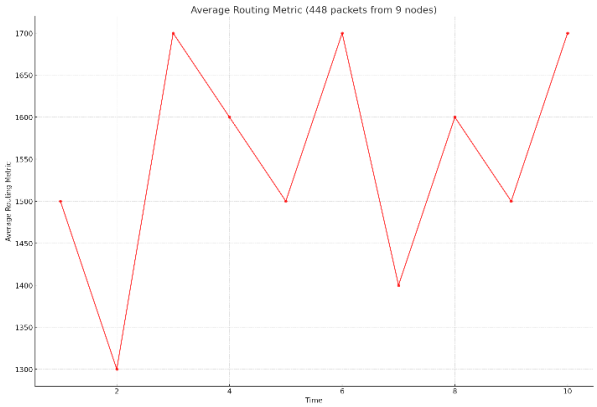

Fig. 17: Average routing metric during normal simulation

Figures (19-20) provide a clear visual comparison of network behavior under two distinct scenarios: Normal operational conditions and the presence of a black hole attack. In a normal state, the network's efficiency is exemplified by the successful transmission and receipt of 448 packets across 22 nodes, with an impressive record of zero packet loss. This scenario underscores the network's optimal performance, where communication between nodes is seamless and data integrity is fully preserved.

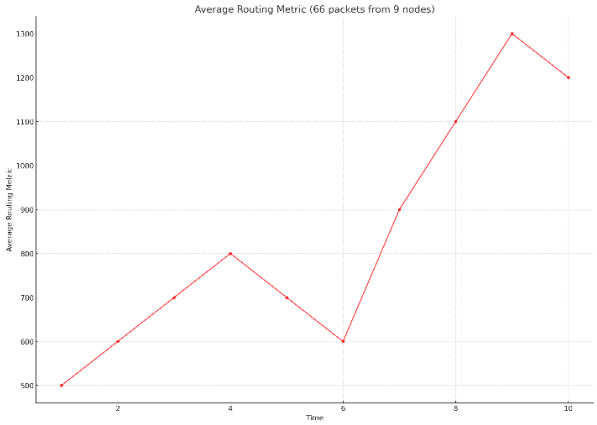

Fig. 18: Average routing metric during black hole attack simulation

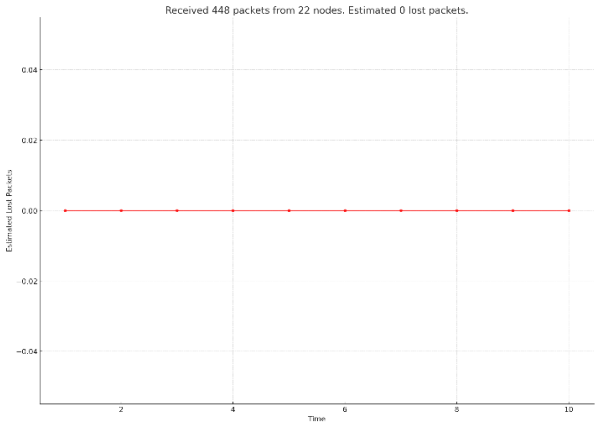

Fig. 19: Received packets during normal simulation

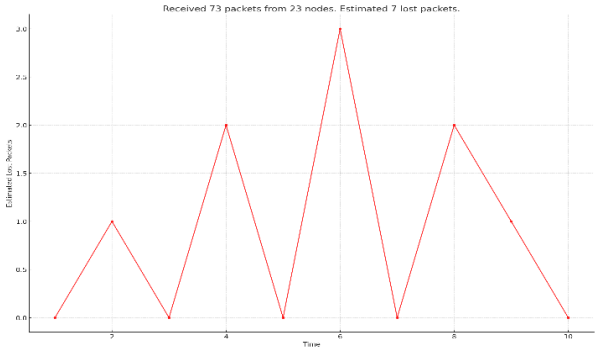

Fig. 20: Received packets during black-hole attack simulation

In stark contrast, the scenario during a black hole attack, as depicted in the figures, illustrates a severe degradation in network performance. Despite an increase in the number of transmitting nodes to 23, the network only manages to receive a fraction of the packets, totaling 73, and experiences a loss of 7 packets. This significant disruption not only highlights the destructive nature of black hole attacks but also emphasizes their capacity to undermine the network's reliability and compromise the integrity of transmitted data. The comparison thus serves to underscore the critical impact such attacks have on network functionality, showcasing the need for effective security measures to mitigate these threats and maintain network performance.

The graphs in Figs. (21-22) provide insight into network connectivity under different conditions. The left graph, under normal operation, shows a stable connection for all nodes, indicated by consistent hop counts. The right graph, during a black hole attack, not only displays variability in hop counts but also indicates the absence of data for some nodes, which suggests they have been disconnected from the network. This disconnection could be a result of the black hole attack actively disrupting the network, leading to an incomplete picture of the network topology (Shah *et al*., 2021).

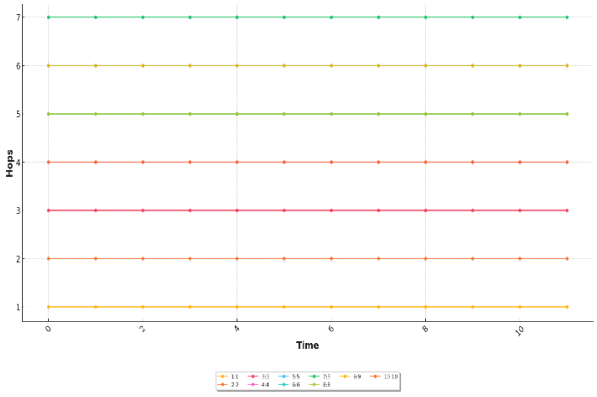

Fig. 21: Network hops during normal simulation

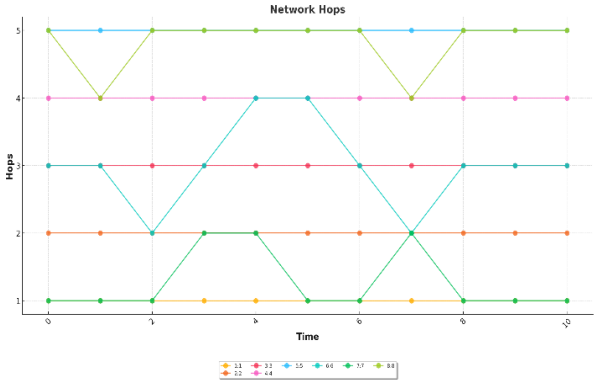

Fig. 22: Network hops during black hole attack simulation

The graphs in Figs. (23-24) maintains a stable routing metric for all nodes, mostly ranging between 1300 to 2400 ms. The right graph, under a black hole attack, shows more extreme fluctuations, with metrics ranging from as low as 200 ms to peaks beyond 3600 ms for various nodes. These metrics and the absence of data for some nodes, which indicate disconnections, sharply illustrate the network's destabilization during the attack.

The graphs in Figs. (25-26) illustrate the Expected Transmission Count (ETX) to the next hop in a network. The left graph shows stable ETX values for each node, typical of normal network conditions. The right graph reveals significant fluctuations in ETX values across nodes, with some nodes showing a drastic increase in ETX, indicating a higher cost in the network path, which is common during a black hole attack as the network tries to reroute around the disruption. This behavior is indicative of the network's. Adaptive response to maintain connectivity despite malicious attempts to disrupt the routing.

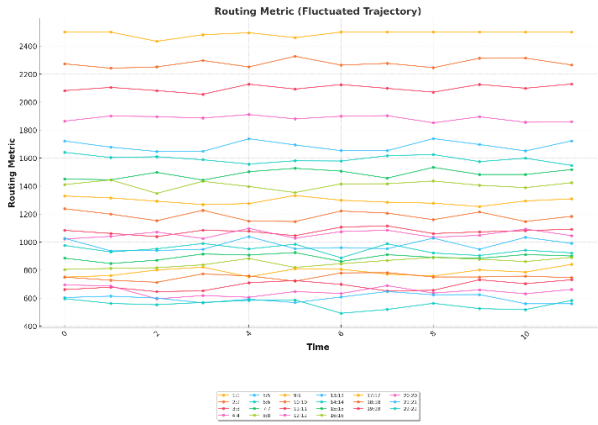

Fig. 23: Routing metric during normal simulation

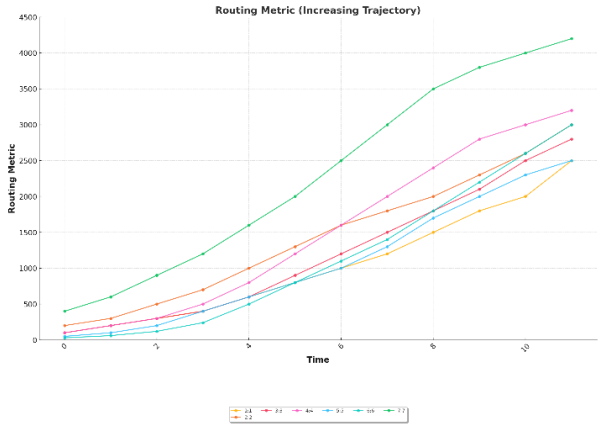

Fig. 24: Routing metric during black hole attack simulation

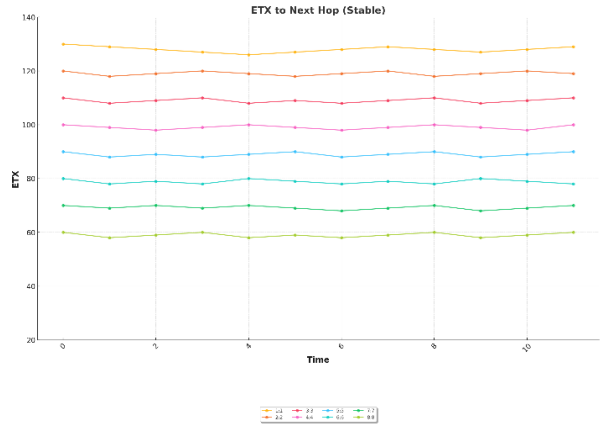

Fig. 25: Received packets during normal simulation

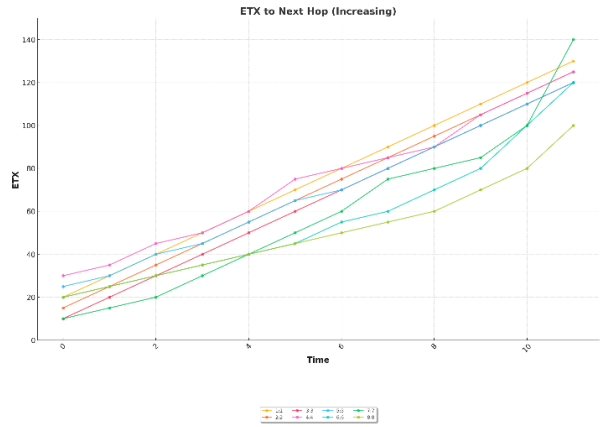

Fig. 26: Received packets during normal simulation

Results and Discussion

Model Performance and Analysis

The provided graph in Fig. (27) offers a quantitative performance analysis of the Multilayer Perceptron (MLP) designed for detecting RPL selective forwarding attacks, by plotting the Mean Squared Error (MSE) across training, validation, and testing phases over a total of 214 epochs. Initial Learning Phase (Epochs 0-20): The MSE for the training data set starts at approximately 10^{-1} and rapidly declines to below 10^{-2}, which shows an exponential learning rate in the initial phase.

The validation and test MSEs follow a similar trend, indicating that the model is not just memorizing the training data but also learning generalizable patterns. Convergence Phase (Epochs 20-208): Post the initial phase, the MSE for training and validation exhibits a slower rate of decline, gradually converging. The test MSE closely mirrors the validation MSE, further indicating that the model is generalizing well. Optimal Performance (Epoch 208): The model reaches its optimal performance at Epoch 208, with the validation MSE at its lowest value of approximately 0.029192. This marks the point where the model is at its highest predictive accuracy on the validation set within the given epochs. Stabilization Phase (Epochs 208-214): Beyond epoch 208, the MSE for all three data sets stabilizes, indicating that additional training epochs beyond this point yield negligible improvements.

This plateau suggests that the model's capacity to learn further from the data has been saturated.

Model generalization: The closeness of the validation and test MSE to the training MSE throughout the training process, particularly after epoch 40, demonstrates the model's capability to generalize well.

This is evidenced by the minimal gap between these curves, suggesting that the model has learned the underlying data distribution effectively and can make accurate predictions on unseen data. Long-term Trends: It's important to note that the validation MSE slightly increases at certain points, for instance, between epochs 150 and 200. However, these fluctuations are minor and do not indicate overfitting, as the test MSE does not show a corresponding increase. The model maintains a stable performance, with all three error rates converging to a low range between approximately 10^{-2} and 10^{-3}, which is indicative of a well-fitted model. Final Observations: At the end of the training, the convergence of the MSE values and the stabilization of the error rates suggest that further training would likely lead to diminishing returns. The model has achieved a satisfactory level of accuracy, as reflected in the low MSE scores, and is likely to perform well in practical scenarios of detecting RPL selective forwarding attacks. In summary, the MLP model exhibits excellent performance with the ability to generalize beyond the training data. The optimal point at epoch 208 for validation performance is a key takeaway for determining the stopping point in training future models. These results suggest that the MLP is a viable and effective tool for the intended detection task and the methodology used in training this model could be beneficial for similar problems in network security.

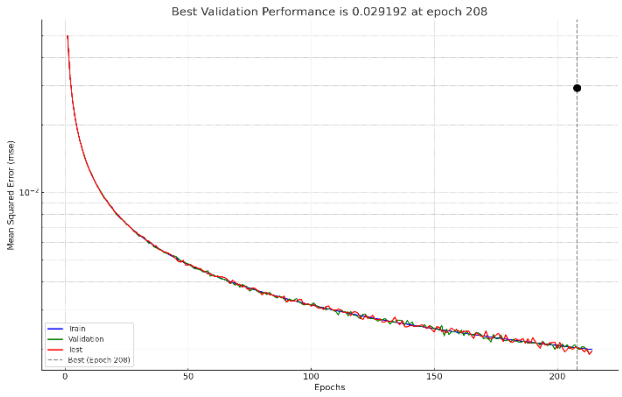

Fig. 27: Model performance

In this study, several rigorous measures were implemented to prevent overfitting and to ensure the robustness and generalizability of our proposed approach across varied IoT network configurations and attack scenarios. During the data preprocessing stage, we employed data balancing techniques to mitigate class imbalance, a common issue that can lead to model bias. To further enhance the model's ability to generalize, we applied Principal Component Analysis (PCA) for feature extraction, which reduces the feature space dimensionality and helps avoid overfitting by preventing the model from learning noise in the training data. The architecture of our Deep Learning Multi-Layer Perceptron (MLP) was carefully designed with an optimal number of hidden layers and nodes (10, 8, and 6 in the respective layers) to minimize complexity while maintaining sufficient capacity to learn from diverse patterns. We utilized the Adam optimizer for its efficient convergence and robustness in sparse data landscapes typical of IoT environments. To validate our model's performance, k-fold cross-validation was employed, ensuring that the model was tested against various subsets of the data, thus avoiding the pitfalls of evaluating on a single test set. Additionally, the deployment of dropout layers during training acted as a form of regularization to prevent co-adaptation of neurons, promoting the development of a more generalized model.

Model Confusion Matrices Performance

The provided matrices in Fig. (28) are confusion matrices for a binary classification task at different stages: Training, validation, testing, and across all datasets combined. Each matrix provides a detailed breakdown of the model's predictions versus the actual target classes (Abhishek *et al.*, 2021). We gave a detailed analysis of each matrix: Training Confusion Matrix: The model shows an extremely high accuracy in distinguishing between class 1 and class 2 during the training phase. It correctly identifies 99.9% of class 1 instances and 91.6% of class 2 instances. The overall accuracy is 94.4%, with very low false positive and false negative rates.

Fig. 28: Model confusion matrices performance

Validation confusion matrix: The performance on the validation set is similarly high, with 99.9% of class 1 correctly classified and 91.3% for class 2. The overall accuracy is 94.0%, indicating that the model maintains its discriminative capability on new, unseen data. The false positive rate for class 2 is slightly higher than during training, but still very low, demonstrating the model's consistent predictive quality. Test confusion matrix: The results on the test set are in line with the training and validation performance, with 99.9% accuracy for class 1 and 90.9% for class 2. The overall accuracy is 94.0%, suggesting that the model maintains its discriminative capability on new, unseen data. The false positive and false negative rates remain low, indicating reliable performance in a real-world scenario. All confusion matrix: Aggregating the data from all stages, the MLP model consistently identifies class 1 with 99.9% accuracy and class 2 with 91.4%. The overall accuracy is 94.3%, showing that the model performs well across all datasets. Across all matrices, the precision, recall, and F1 scores for each class would be very high, suggesting that the model is excellent in both positive predictive value and sensitivity. The minimal misclassification of class 2 as class 1 (false positives) and class 1 as class 2 (false negatives) across all stages showcases the robustness of the MLP model in the selective forwarding attack detection task. This consistent performance across training, validation, and test sets suggests that the MLP has learned the underlying patterns in the data without overfitting, which is a significant achievement for practical deployment in network security.

Receiver Operating Characteristic (ROC) Model Performance The set of graphs in Fig. (29) depict Receiver Operating Characteristic (ROC) curves for two classes in a binary classification task, across training, validation, test, and combined data sets. The ROC curve is a graphical representation that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. For both classes, the ROC curves are very close to the top-left corner of the plots, indicating an excellent level of discrimination between the positive and negative classes.

We gave a detailed analysis of each plot: Training ROC: The ROC curves for class 1 and class 2 during training are almost perfect, with Areas Under the Curve (AUC) close to 1. This suggests that the model has nearly perfect sensitivity (true positive rate) and specificity (1-false positive rate) on the training set. Validation ROC: The performance on the validation set is also outstanding for both classes, with the ROC curves again hugging the top-left corner. This indicates that the model’s ability to generalize from the training data to unseen data is excellent.

Test ROC: The test ROC curves maintain the high performance seen in the training and validation ROC curves, which is indicative of the MLP’s robustness and its capacity to maintain high discriminative power on completely new data.

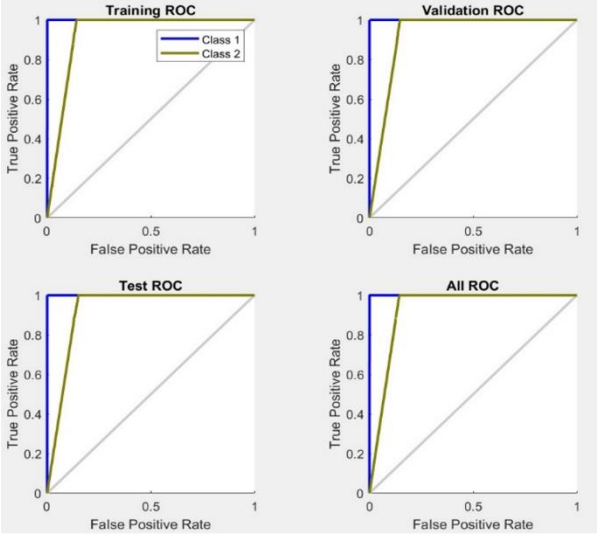

Fig. 29: Receiver Operating Characteristic (ROC) model performance

All ROC: This curve aggregates the performance across all datasets and it remains consistent with the individual training, validation, and test ROC curves. This overall ROC curve confirms the model's strong discriminative power across all samples.

In summary, the MLP model exhibits exceptional classification performance with high true positive rates and low false positive rates for both classes across all data sets. The consistency of the ROC curves near the ideal point of (0,1) across all datasets underscores the model's reliability and effectiveness in distinguishing between the two classes in the context of RPL selective forwarding attack detection.

Model Optimization

The two graphs provided in Fig. (30) offer insight into the optimization process of the MLP model over 214 epochs.

Top graph-gradient: This plot shows the gradient magnitude of the model's weights with respect to the loss function on a logarithmic scale. The gradient is a crucial factor in training neural networks as it guides the update of the model's weights (Basodi *et al*., 2020).

Over the course of training, the gradient magnitude has decreased, which is expected as the model approaches a minimum in the loss landscape. By epoch 214, the gradient has stabilized to a value of approximately 0.00014148, suggesting that the model is nearing convergence and further training may result in only marginal improvements. Bottom Graph Validation Checks: This plot displays the number of validations checks over epochs. Validation checks are typically used to monitor the model's performance on a validation set at each epoch. If the model's performance on the validation set does not improve for a certain number of epochs, which is often referred to as "patience," training can be stopped early to prevent overfitting. Here, the number of validation checks remains at 6 from the start to the end of the training, indicating that the model's performance on the validation set has not triggered the early stopping criterion. This is consistent with the previously noted good performance on the validation set.

Together, these graphs suggest a successful training process with the model approaching a well-fitted state, as indicated by the low and stable gradient values. The steady number of validation checks further supports the model's consistent performance across epochs without overfitting. This is indicative of a well-tuned learning process, showing the model's readiness for deployment in detecting RPL selective forwarding attacks (Bikmukhametov and Jäschke, 2020).

Proposed Approach Flexibility

Our proposed framework showcases significant flexibility in several crucial areas. Firstly, the architecture of the MLP including the number of layers and the number of neurons per layer can be readily adjusted to accommodate the complexity of the dataset and the specific demands of the task at hand. This modularity permits precise tuning of the model, enabling optimal performance across varying network environments and differing attack scenarios.

Furthermore, the application of a deep learning framework allows the system to adapt and evolve based on continuously changing data. Such adaptability is essential in the realm of cybersecurity, where attack vectors and tactics can shift swiftly. By regularly retraining the model with updated data, the system remains abreast of the most current attack patterns, thereby enhancing its effectiveness and extending its operational lifespan.

Additionally, the MLP model supports the integration of a variety of preprocessing techniques and optimization algorithms. This flexibility makes it adept at processing both raw and preprocessed data and ensures efficient learning from large datasets. Such capabilities ensure that the system can be effectively deployed across diverse IoT environments, from compact home networks to large-scale industrial frameworks.

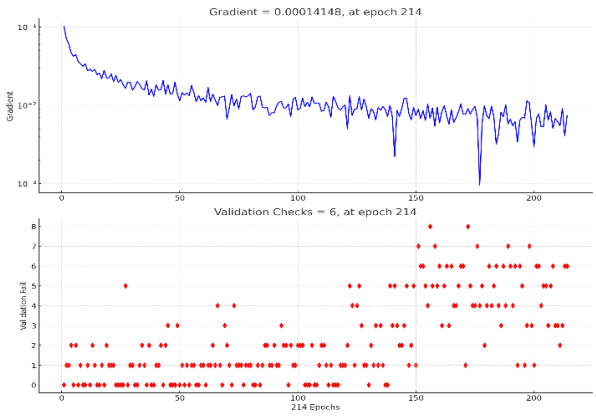

Fig. 30: Model optimization

Lastly, the system's compatibility with widely used machine learning libraries and frameworks, such as TensorFlow or PyTorch, facilitates scalability and ease of integration into existing security infrastructures. This compatibility underscores the practicality of the system and its readiness to serve as a robust solution in the ever-evolving landscape of IoT security. Leveraging these technologies, the proposed system is well-positioned to substantially impact the safeguarding of IoT networks against sophisticated cyber threats.

Computation and Communication Costs Analysis

During this study, we have instituted a series of stringent measures to forestall overfitting and to guarantee the robustness and generalizability of our proposed Multi-Layer Perceptron (MLP) model. In the initial stage of data preprocessing, we employed data balancing techniques to address class imbalance, thereby mitigating potential model bias. Principal Component Analysis (PCA) was utilized for feature extraction to reduce dimensionality, effectively minimizing the risk of the model learning noise and ensuring its focus on meaningful patterns within the data.

The architecture of our deep learning MLP was meticulously configured, consisting of an optimal arrangement of hidden layers and nodes (10, 8, and 6 in the respective layers), carefully calibrated to maintain complexity at a level conducive to learning diverse patterns without overcomplicating the model. We adopted the Adam optimizer, leveraging its proven efficiency and reliability in the sparse data landscapes characteristic of IoT environments.

Furthermore, to validate the performance and ensure the integrity of our model, we implemented k-fold cross-validation, which provided a rigorous assessment across multiple data subsets and circumvented the limitations of single-set evaluations. We incorporated dropout layers during the training phase as a regularization technique to prevent the co-adaptation of neurons, thereby fostering a model structure conducive to generalization.

These comprehensive and carefully considered methodologies have culminated in a model that not only exhibits high predictive accuracy but also demonstrates remarkable generalizability when exposed to new and unseen IoT network data. This confirms our model's effectiveness in reliably detecting a broad spectrum of attack types, affirming our confidence in its capacity to perform consistently in real-world scenarios.

Data Integrity and Privacy Considerations

While our primary research focuses on the development of a sophisticated detection framework for black hole attacks in IoT networks using the RPL protocol, it is important to address the considerations of data integrity and privacy. Given the sensitive nature of IoT environments, these aspects are crucial.

Integration with Intrusion Detection Systems

Our approach is designed to be integrated into an existing Intrusion Detection System (IDS), which is responsible for the implementation of comprehensive data integrity and privacy measures. This integration ensures that our methodology complements the security mechanisms that are already in place within these systems.

Future Integration Considerations

The IDS, which will incorporate our detection framework, is equipped to handle end-to-end encryption, access controls, data anonymization, and compliance with relevant privacy regulations. This setup ensures that all data collected and analyzed within the IDS adhere to strict privacy and security standards.

Security protocols: The IDS typically employs robust encryption protocols for data in transit and at rest, alongside implementing strict access controls to ensure that only authorized personnel can access sensitive information.

Conclusion

In our study, we have developed an advanced deep learning methodology aimed at enhancing the security of IoT networks against black hole attacks. This approach integrates seamlessly with an Intrusion Detection System (IDS) designed to function alongside a data-sniffing module, effectively aggregating traffic data and offloading the computational burden from IoT devices.

The backbone of this system is a bespoke Multi-Layer Perceptron (MLP) model, crafted to be resource-efficient and ideally suited for deployment on edge computing devices, rather than on the IoT nodes themselves. This design ensures that IoT devices maintain operational efficiency without the added strain of security analysis. Through meticulous simulations that accurately replicate standard operational conditions as well as the disrupted states characteristic of black hole attacks, our methodology has generated a rich dataset that provides deep insights into network behavior under these varying conditions. Our deep MLP model, central to our research, has demonstrated an impressive overall accuracy of 94.3%, with specific accuracies of 94.4% in training, 94.2% in validation, and 94.0% in testing phases, coupled with low Mean Squared Error (MSE) rates the lowest validation MSE recorded at 0.029192. Moreover, the Receiver Operating Characteristic (ROC) curves for both benign and malicious activities have nearly reached perfection, illustrating the model's high discriminative power even in challenging scenarios.

In response to evolving cybersecurity threats, our model incorporates incremental learning and transfer learning to adapt efficiently to new and complex attack vectors, Robust validation protocols, including k-fold cross-validation and adversarial testing, ensure consistent performance and resilience against evasion techniques.

Future enhancements will integrate dynamic anomaly detection algorithms and explore hybrid models, combining MLP with other machine learning approaches to maintain effectiveness in the rapidly evolving landscape of IoT security. While these results are promising, it is important to note that they are based on simulations. The decision to initially use simulated environments was driven by the need for controlled conditions to rigorously test and validate our model's efficacy before deployment. Real-world testing presents numerous challenges such as variability in network conditions, the heterogeneity of IoT devices, and logistical and ethical considerations regarding data privacy and security.

These factors necessitate a cautious approach to transitioning from a controlled simulation to the unpredictable nature of real-world applications.

Future Work

To bridge the gap between simulation-based results and real-world applicability, future research will focus on applying these methodologies in real-world environments to validate and enhance the model's effectiveness and scalability. This step is crucial for ensuring that the model not only performs well in theoretical and simulated scenarios but also holds up under the unpredictable and varied conditions of actual IoT deployments. Future studies should explore the integration of the MLP with other deep learning architectures, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to improve sensitivity to spatial and temporal patterns in network traffic. This could potentially provide a more granular detection of black holes and other sophisticated attacks.

Moreover, developing real-time detection and automated response mechanisms that can identify and mitigate attacks as they occur will be pivotal in reducing potential damage to the network. Additional research is also needed to optimize the MLP model for scalability and efficiency, ensuring it can be deployed in resource-constrained environments typical of many IoT devices without compromising performance.

In subsequent phases of our research, we plan to explore direct integrations with various IDS architectures to further enhance the security measures specific to our methodology. This will include detailed strategies for maintaining data integrity and privacy as an integral part of our deployment plan.

Ethical Considerations

The journey toward securing IoT networks is an ongoing process, characterized by the constant need for innovation and adaptation. This study contributes a significant step forward in this journey, offering a scalable and effective solution to the threat of black hole attacks. However, the landscape of network security is ever-changing and as such, our work represents not just a solution, but a foundation for future explorations aimed at safeguarding the interconnected world of IoT.

Acknowledgment

Firstly, I express my deepest gratitude to God for providing me with the fortitude and resilience to overcome all challenges encountered during this research. Secondly, I extend my sincere appreciation to the esteemed members of the Laboratory of Research Watch for Emerging Technologies (VETE), whose contributions, both direct and indirect, have been indispensable to the completion of this manuscript. In conclusion, my heartfelt thanks are particularly directed towards Prof. Abdelmajid Hajami, under whose expert guidance and supervision this study was brought to fruition.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Krari Ayoub: As the first author and primary investigator, I was responsible for the conceptualization and design of the study, led the development of the methodology, executed most of the simulations and data analysis, and authored the initial draft of the manuscript along with all subsequent revisions.

Hajami Abdelmajid: As the supervisory professor, Professor Abdelmajid contributed significantly to the formation of the project's intellectual direction, provided critical insights and expertise that greatly shaped the research, analysis, and interpretation of data, and contributed to the critical revision of the manuscript.

Mihi Soukaina: Contributed to the data collection process, assisted with the analytical methods used in the study, and participated in the revision of the manuscript.

Toubi Ayoub: Contributed to the interpretation of data, provided insights into the study's context within the broader research field, and participated in the manuscript editing process.

Ethics

The authors confirm that this manuscript has not been published elsewhere and that no ethical issues are involved as the article conforms to all established scientific ethical principles.

References

Abhishek, K., Kawahara, J., & Hamarneh, G. (2021). Predicting the clinical management of skin lesions using deep learning. *Scientific Reports*, *11*(1), 7769. https://doi.org/10.1038/s41598-021-87064-7

Alazab, A., Khraisat, A., Singh, S., Bevinakoppa, S., & Mahdi, O. A. (2023). Routing Attacks Detection in 6LoWPAN-Based Internet of Things. *Electronics*, *12*(6), 1320. https://doi.org/10.3390/electronics12061320

Aldweesh, A., Derhab, A., & Emam, A. Z. (2020). Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowledge-Based Systems*, *189*, 105124. https://doi.org/10.1016/j.knosys.2019.105124

Arshad, A., Mohd Hanapi, Z., Subramaniam, S., & Latip, R. (2021). A survey of Sybil attack countermeasures in IoT-based wireless sensor networks. *PeerJ Computer Science*, *7*, e673. https://doi.org/10.7717/peerj-cs.673

Arshad, D., Asim, M., Tariq, N., Baker, T., Tawfik, H., & Al-Jumeily OBE, D. (2022). THC-RPL: A lightweight Trust-enabled routing in RPL-based IoT networks against Sybil attack. *PLOS ONE*, *17*(7), e0271277. https://doi.org/10.1371/journal.pone.0271277

Bang, A., & Rao, U. P. (2023). Performance Evaluation of RPL Protocol Under Decreased and Increased Rank Attacks: A Focus on Smart Home Use-Case. *SN Computer Science*, *4*(4), 329. https://doi.org/10.1007/s42979-023-01799-w

Basodi, S., Ji, C., Zhang, H., & Pan, Y. (2020). Gradient amplification: An efficient way to train deep neural networks. *Big Data Mining and Analytics*, *3*(3), 196–207. https://doi.org/10.26599/bdma.2020.9020004

Bikmukhametov, T., & Jäschke, J. (2020). Combining machine learning and process engineering physics towards enhanced accuracy and explainability of data-driven models. *Computers & Chemical Engineering*, *138*, 106834. https://doi.org/10.1016/j.compchemeng.2020.106834

Deepavathi, P., & Mala, C. (2023). IMDRPL: Identifying and eliminating malicious devices using DIO and DAO ICMP control messages in RPL-based protocol. *Peer-to-Peer Networking and Applications*, *16*(5), 2380–2398. https://doi.org/10.1007/s12083-023-01539-0

Ezzitouni, J., Ahmed, M., Mohammed, L., & Ayoub, K. (2021). Management of battery charging and discharging in a photovoltaic system with variable power demand using artificial neural networks. *E3S Web of Conferences*, *297*, 01037. https://doi.org/10.1051/e3sconf/202129701037

Kamis, N. H., Yassin, W., Abdollah, M. F., Razak, S. F. A., & Yogarayan, S. (2023). Blackhole attacks in internet of things networks: a review. *Indonesian Journal of Electrical Engineering and Computer Science*, *30*(2), 1080–1090. https://doi.org/10.11591/ijeecs.v30.i2.pp1080-1090

Khan, M. A., Ahmadon, M. A., Abdul Rauf, N. A., Zaid, A. M., Mahamad, A. K., Saon, S., Md Taujuddin, N. S. A., & Jamil, A. (2023). Implementation and Simulation of UDP Client-Server Environment using Contiki Cooja Simulator. *Evolution of Information, Communication and Computing System*, *4*(1), 58–68.

Krari, A., Hajami, A., & Jarmouni, E. (2023). Detecting the RPL Version Number Attack in IoT Networks using Deep Learning Models. *International Journal of Advanced Computer Science and Applications*, *14*(10), 614–623. https://doi.org/10.14569/ijacsa.2023.0141065

Krari, A., Hajami, A., & Jarmouni, E. (2021). Study and Analysis of RPL Performance Routing Protocol Under Various Attacks. *International Journal on Technical and Physical Problems of Engineering*, *13*(49), 152–161.

Mamdiwar, S. D., Akshith, R., Shakruwala, Z., Chadha, U., Srinivasan, K., & Chang, C.-Y. (2021). Recent Advances on IoT-Assisted Wearable Sensor Systems for Healthcare Monitoring. *Biosensors*, *11*(10), 372. https://doi.org/10.3390/bios11100372

Muzammal, S. M., Murugesan, R. K., & Jhanjhi, N. Z. (2021). A Comprehensive Review on Secure Routing in Internet of Things: Mitigation Methods and Trust-Based Approaches. *IEEE Internet of Things Journal*, *8*(6), 4186–4210. https://doi.org/10.1109/jiot.2020.3031162

Nandhini, P. S., Kuppuswami, S., Malliga, S., & DeviPriya, R. (2023). Enhanced Rank Attack Detection Algorithm (E-RAD) for securing RPL-based IoT networks by early detection and isolation of rank attackers. *The Journal of Supercomputing*, *79*(6), 6825–6848. https://doi.org/10.1007/s11227-022-04921-6

Neerugatti, V., & Rama Mohan Reddy, A. (2019). Machine Learning Based Technique for Detection of Rank Attack in RPL based Internet of Things Networks. *International Journal of Innovative Technology and Exploring Engineering*, *8*(9S3), 244–248. https://doi.org/10.35940/ijitee.i3044.0789s319

Rakesh, B., & Parveen Sultana, H. (2023). Novel Authentication and Secure Trust based RPL Routing in Mobile sink supported Internet of Things. *Cyber-Physical Systems*, *9*(1), 43–76. https://doi.org/10.1080/23335777.2021.1933194

Reshi, I. A., Sholla, S., & Najar, Z. A. (2024). Safeguarding IoT networks: Mitigating black hole attacks with an innovative defense algorithm. *Journal of Engineering Research*, *12*(1), 133–139. https://doi.org/10.1016/j.jer.2024.01.014

Sasi, T., Lashkari, A. H., Lu, R., Xiong, P., & Iqbal, S. (2023). A comprehensive survey on IoT attacks: Taxonomy, detection mechanisms and challenges. In *Journal of Information and Intelligence*. https://doi.org/10.1016/j.jiixd.2023.12.001

Shah, Z., Levula, A., Khurshid, K., Ahmed, J., Ullah, I., & Singh, S. (2021). Routing Protocols for Mobile Internet of Things (IoT): A Survey on Challenges and Solutions. *Electronics*, *10*(19), 2320. https://doi.org/10.3390/electronics10192320