

Impact of Machine Learning-Based Security Mechanisms on Quality of Service in Internet of Things Networks

Sheetal Jariya¹, Mohit Jain²
Research Scholar¹, HOD CSE²

sheetaljariya2214@gmail.com¹, bmctmohitcs@gmail.com²

Abstract: *The emergence of the Internet of Things (IoT) and other innovative technologies has enabled the interconnection of devices worldwide, earning them the moniker "smart gadgets" due to their capabilities to send, receive, and process data. This technology is experiencing rapid growth, with a continually increasing user base. The success of IoT hinges on factors such as data transmission rates, quality of service (QoS) maintenance, and management of energy constraints in battery-operated devices. At the network level, QoS is evaluated based on metrics including end-to-end delay, throughput, jitter, and packet delivery ratio. With the proliferation of IoT devices, ensuring both device and data security in network communications becomes paramount. This paper delves into algorithms employed to safeguard the locations of source and sink nodes from potential breaches. Additionally, it investigates the impact of AODV protocols on the QoS offered by IoT networks. Malware poses significant threats in this context, prompting researchers, industry professionals, and end-users to seek effective countermeasures. Early and accurate prediction of malware behavior is crucial to mitigate potential damage. The research endeavors to combat malware using the K-Nearest Neighbors (KNN) algorithm, predicting its behavior and eliminating it. Utilizing these classifiers appropriately can significantly enhance prediction accuracy.*

Keywords: *K-Nearest Neighbor, IoT, QoS, Malware detection, Wireless Network, Security.*

1. INTRODUCTION

The Internet of Things (IoT) is a system that lets physical objects be linked together and monitored over the internet. Things like computers, digital machines, electrical or home appliances, and so on, all of which have their own digital identities, are connected to the things around them and can share data with them. This makes it possible for the objects to connect and talk to each other in a smart way. With the help of the Internet of Things (IoT), things can connect without any interaction between people or between people and digital devices. Even though we think of tablets, laptops, computers, and cell phones as ways to connect, in reality, things can connect without us having to do anything. The needs of people who use the Internet of Things have led service providers to make a wide range of apps. The quality of services (QoS) that customers want from an app can vary from person to person. Similarly, the QoS of the many apps that use the internet of things will also be different (IoT)[1].

For each application, the quality metrics should be set in a very clear way, so that a user can tell the service provider what he expects and the service provider can make changes to meet those expectations. So, the researchers should put most of their efforts into finding the QoS (Quality of Service) indicators to find out what IoT service users want. Due to how quickly the internet has grown, the number of cyber risks caused by malware has also gone up. One definition of malware says that it is a type of computer programme that is made to hurt the other user's computer in a number of ways. Malware comes in so many different forms now, and anyone can buy malware on the dark web to increase the number of attacks they launch against our system. This makes it very hard for anti-virus software to fully protect a computer. Malware, also called "malicious software," is a programme that sneaks into a computer system without the user's permission and tries to damage the system or steal private information that is stored on the system. Malicious software, which is also called "malware," is any piece of software that

is made to do something bad on purpose by an enemy. Malwares are called things like viruses, worms, Trojan Horses, root-kits, spyware, backdoors, botnets, and adware, among other things, based on how they behave and how they infect computers. [2] Every day, thousands of new malicious software programmes are made, and the structure of already existing malicious software programmes is always changing, making it harder and harder to find them[3].

Symantec's most recent report on internet threats says that 317 million new types of malicious software have been found. Because more samples of malware are being made every day, automated tools and methods are needed to tell the difference between harmful and harmless code. Signature-based malware classification is used by almost all anti-virus software on the market[4]. This method compares the unknown malware to a database of known malicious programmes to find out if the file in question has malware or is safe to use. A unique identifier that can be added to a binary file is called a signature. Malware's signature can be found through static analysis, dynamic analysis, or a combination of the two. Once the signature is found, it is saved in a database called the signature. The biggest problem with this strategy is that the signature database needs to be updated often because new malware is created so quickly every day. Symantec's latest report on internet threats says that 317 million new types of malware have been found. Because there are more new samples every day, automated tools and methods are needed to tell the difference between malicious and safe code. Most commercial anti-virus software uses a method based on signatures to sort malware. This method compares unknown malware to a database of known malware to figure out if a file is malware or not. The signature is a way to identify a binary file in a unique way. Malware's signature can be found using static analysis, dynamic analysis, or a mix of the two. The signature is then stored in a database called a "signature database." The biggest problem with this method is that the signature database needs to be updated often because new malware comes out every day [5-6].

IoT helps connecting help physical world to connect to computer world. As its application are increasing day by day, privacy issues are also increasing. Different attacks like spoofing, DDoS attack, and jamming, malware and eavesdropping are becoming potential threats. Small IoT devices are restricted to execute computational-intensive and latency sensitive security tasks. Today, the IoT devices are protected using authentication, in which source nodes are identified and identity based attacks are prevented, access control, secure offloading techniques and malware detection to prevent against privacy leakages. These techniques are not

applied to small IoT devices like outdoor sensors, so spoofing attack on them is not recognized [7]. Machine learning techniques are applicable on IoT devices whether small or large. These techniques include: Supervised Learning: This includes support vector machine, naïve Bayes, neural network, deep neural network, random forest, K nearest neighbor to track network traffic or app traces of IoT devices to build classification or regression models [8].

2. LITERATURE SURVEY

AODV The author of [9] says that the main goals of WSNs and IoT are to reduce the amount of power used to make the network last longer and to make sure it is safe. Mamdani An energy efficient secure route adjustment (ESRA) model, which is explained in [10], used fuzzy logic to figure out the most energy-efficient way to communicate. It figures out the best route by adding up the values of a number of quality-of-service criteria[11]. Sink nodes can be moved, but to set up a new route, you need to know where the currently used sink node is. This method uses little power because it chooses the route based on how reliable it is. Cluster-based routing algorithms need a lot more energy to work because there are so many intermediate nodes in the path of communication[12].

The authors of [13] came up with a double level unequal clustering algorithm (DLUC) to solve the problem of clustering techniques using more power than they used to. This algorithm tells each cluster how much traffic it needs to handle. Since the cluster heads don't have to be present for information to flow between nodes, the number of clusters and nodes along the transmission line can be cut [14]. Networks can use less energy if the bandwidth is optimized, the values of interference between control packets are lowered by using framing periods to avoid congestion, and the values of data loss are lowered. This method doesn't take into account how people move around and how different clusters are sized, which are two major factors that cause networks to use more power [15].

This section will try to list a few ways that have been made to keep track of the positions of source nodes, sink nodes, or sometimes both source and sink nodes, while an Internet of Things application is running [16]. We also looked into how security algorithms affect how efficiently they use energy and how they affect quality of service measures like throughput, end-to-end delay, and packet delivery ratio. A. Where the source node is the idea of source location privacy (SLP) has become a difficult problem for research institutions to solve if they want to keep their networks safe [17]. If SLP is not present, it will be easier to

figure out where the source nodes are and get to the data before it is sent along the communication route [18]. Most of the time, a route made by the sensors will have a source node, a sink node, and a few nodes in between. In order to move data packets, these intermediate nodes will use hopping techniques. Some of the research shows that it is easy for the source's enemies to figure out where the source is, even if they know where some of the nodes are along the route that is currently being taken [19]. Because of this, it is very important to make SLP algorithms to protect IoT networks from security attacks. In [20] a technique called SDR-m was described. SDR-m is a stochastic and diffuse routing.

3. PROPOSED METHODOLOGY

The workflow begins with system initialization, where the simulation environment is configured by defining essential network parameters such as the number of nodes, transmission range, routing protocol, simulation duration, energy model, and communication settings. Following initialization, nodes are deployed randomly or according to a predefined topology within the network area. Each node is assigned a unique identifier, initial energy level, communication range, and location, enabling the formation of the network topology through neighbor discovery. Once the network becomes operational, the system continuously monitors node activities to detect attack nodes. Malicious nodes are identified by analyzing abnormal behaviors such as packet dropping, excessive delays, route manipulation, abnormal traffic generation, or unauthorized forwarding using predefined detection algorithms or machine learning techniques. After detecting malicious nodes, the compromised communication links are removed, and the routing protocol reroutes the data through an alternative secure path that excludes the identified attackers. This rerouting process ensures reliable packet transmission while maintaining network security and communication continuity. Finally, the proposed framework undergoes performance evaluation, where its effectiveness is assessed using Quality of Service (QoS) and security metrics, including packet delivery ratio, throughput, end-to-end delay, energy consumption, network lifetime, detection accuracy, false positive rate, and routing overhead. These performance measures demonstrate the capability of the proposed approach to provide secure and efficient communication in IoT networks.

4. ANALYSIS AND RESULT DISCUSSIONS

In After reviewing various papers, it became apparent that routing algorithms can effectively maintain certain quality of

service (QoS) parameters such as end-to-end delay, throughput, and packet delivery ratio. However, many location privacy algorithms struggle to strike a balance between energy efficiency and security. This insight was gleaned from the comprehensive review of relevant literature. While numerous studies have succeeded in enhancing the security of source or sink nodes, the introduction of fraudulent packets along the route tends to increase energy consumption. There exists a direct and inverse relationship between network energy consumption and its longevity. As the energy demands of privacy algorithms rise, the lifespan of the network diminishes. This presents a significant challenge in implementing security and privacy algorithms for Internet of Things (IoT) devices with limited energy resources.

To mitigate the need for frequent route recovery, a backup route mechanism can be established within the coverage area of a Roadside Unit (RSU) that has recently passed by. Implementing this approach can reduce the frequency of route rediscovery. The optimal method involves creating a backup route from a serving RSU that has already traversed the area, thus enhancing network reliability and efficiency while conserving energy resources.

Case -1- The scalability of a VANET, irrespective of the number of nodes, is not contingent on its ability to grow limitlessly. VANETs facilitate both Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications, where nodes acquire information from other nodes or Roadside Units (RSUs), necessitating accurate data transmission. Security requirements differ across VANETs, especially regarding inter-vehicle communication.

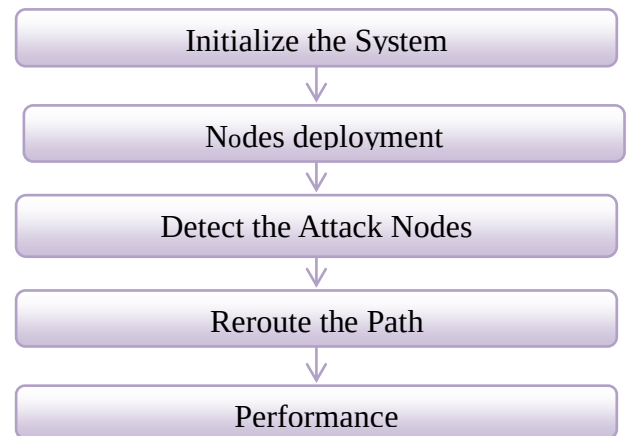


Figure 1:- Simulation Flow Diagram

Figures 2 through 6 illustrate various aspects of network initialization, node-to-node data transfer, AODV paths corresponding to changes in vehicle position, delay for

different data rates, energy consumption across different data rates, the number of paths found during simulation, and network lifetime across different data rates, respectively. Figure 2 depicts the initialization of the network, while Figure 3 illustrates the process of node-to-node data transfer. Figure 4 showcases AODV paths structured based on changes in vehicle position, and Figure 5 presents the delay experienced for different data rates. Energy consumption

across different data rates is depicted in Figure 6, and Figure 7 indicates the number of paths discovered during simulation. Finally, illustrates the network lifetime across various data rates. These figures collectively provide a comprehensive understanding of the network dynamics and performance metrics associated with different data rates, facilitating insights into network efficiency, reliability, and longevity.

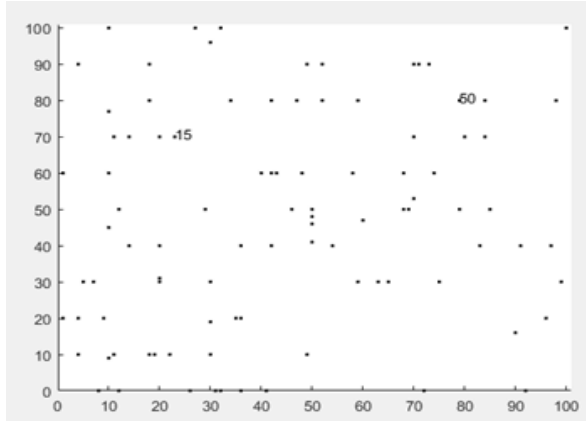


Figure 2:- Initialization Network

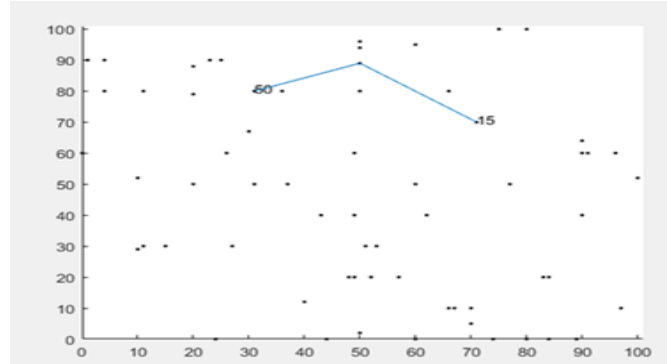


Figure 3:- Node To Node Data Transfer

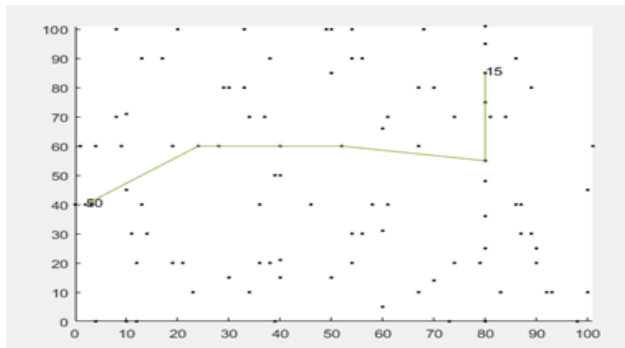


Figure 4:- AODV Paths for Each Change in Vehicle Position Into a Structure

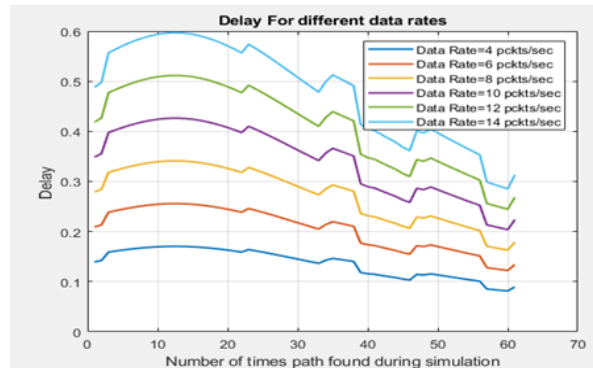


Fig.5 Delay for Different Data Rates

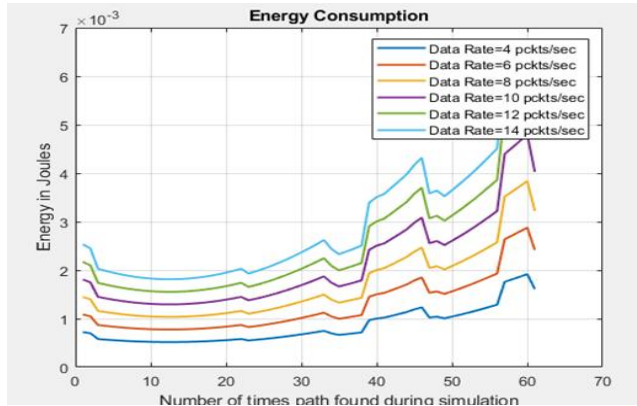


Figure 6:- Energy Consumption for Different Data Rates

Case-2

- The proposed system aims to enhance the quality of service (QoS) and security of wireless networks by employing adaptive measures triggered when performance evaluation metrics fall below predefined thresholds. A key aspect of this enhancement involves utilizing the K Nearest Neighbors (KNN) algorithm for malware detection.

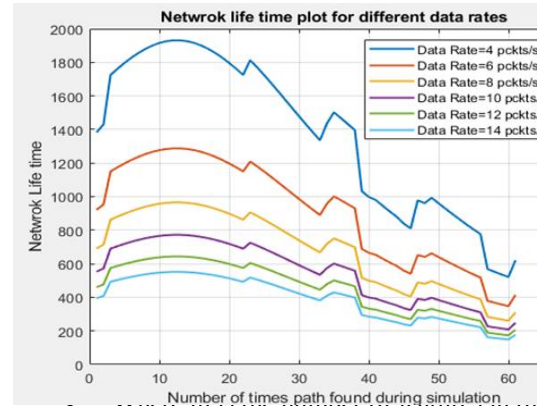


Fig.7 Network Life Time for Different Data Rates

- Once the nearest neighbors are identified, the majority class among these neighbors is assigned to the query point x_q this process can be represented by the following mathematical expression:
 - $\hat{y}_q = \operatorname{argmax}_{y_i} \sum_{j=1}^m (y_i - y_j)^2$
 - Where $\hat{y}_q$ the predicted class label for the query point is x_q is the indicator function, and k is the number of neighbors.

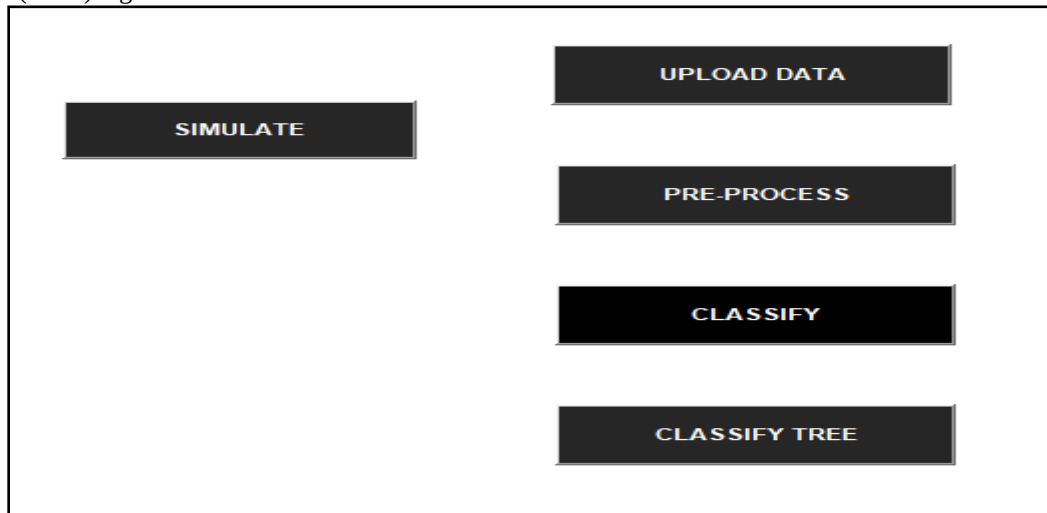


Figure 8:- Flow Diagram for Malware Detection

To ensure the security of sink nodes in wireless sensor networks (WSNs), algorithms are required to conceal their locations and prevent attackers from accessing data packets destined for these nodes. One approach to achieving this is through the use of K-means cluster-based methods,

K-means Clustering for Source and Sink Node Location: The K-means clustering algorithm is applied to track the locations of both source and destination (sink) nodes in WSNs. Fake nodes are introduced alongside real source and sink nodes to enhance security. Data packets can be routed to multiple sink nodes through the clustering

mechanism, maintaining the routing path's length while safeguarding data flow. Real packets are sent through the shortest route, reducing latency and enhancing safety time.

Malware Detection Using K-Nearest Neighbors (KNN):

Supervised learning methods, such as KNN, can be employed for malware detection in IoT devices by analyzing application behavior. The KNN method categorizes network traffic by assigning it to the category with the most items among its K nearest neighbors.

Data Collection and Pre-processing: Malware dataset derived from open-source tools undergoes pre-processing for efficient training.

Cleaning: Removal of irrelevant entries and assigning integer values to non-relevant data.

Transformation: Conversion of non-integer values to integers for computational suitability.

Reduction: Removal of irrelevant features from the dataset that do not contribute to predicting malware categories.

K-means Clustering: K-means clustering involves minimizing the sum of squared distances between data points and their respective cluster centroids. This can be mathematically represented as: Minimize

$$\sum_{i=1}^k \sum_{x \in C_i} \|x - u_i\|$$

Where:

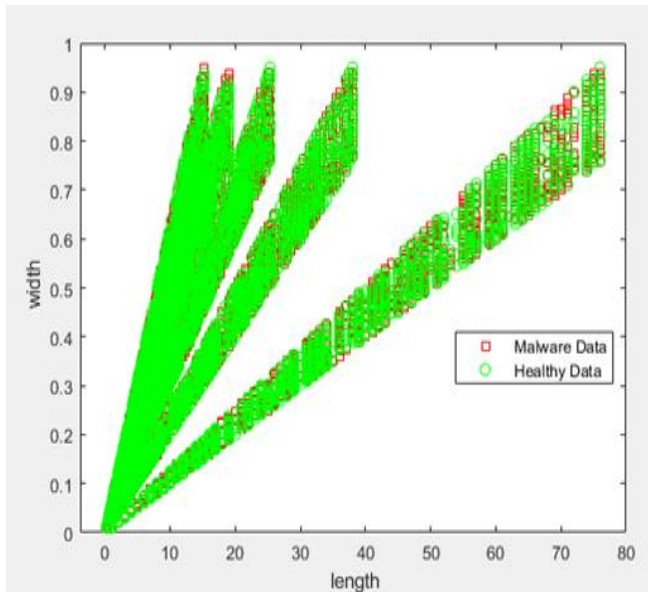


Figure 9:- training of malware and healthy data

C_i represents the i th cluster. u_i represents the centroid of cluster C_i

KNN for Malware Detection:

The KNN algorithm computes the distance between a query point and all other points in the dataset, then assigns the query point to the majority class among its K nearest neighbors. this can be expressed as:

$$\hat{y}_q = \operatorname{argmax}_{y_i} \sum_{j=1}^k I(y_i = y_j)$$

Where:

$\hat{y}_q$ is the predicted class label for the query point.

I is the indicator function.

k is the number of neighbors.

Figures 9 through 12 provide a detailed view of the malware detection process using machine learning. Figure 9 illustrates the training phase where both malware and healthy data are used to train the model. Figure 10 shows the observation phase where the characteristics of healthy data and malware data are analyzed in fig 11. depicts the deletion process of identified malware data, ensuring that only non-malicious data remains. Finally, Figure 12 demonstrates the separation of malware from healthy data, highlighting the effectiveness of the model in distinguishing between the two. These figures collectively showcase the steps involved in detecting, analyzing, and eliminating malware to maintain the integrity of the network.

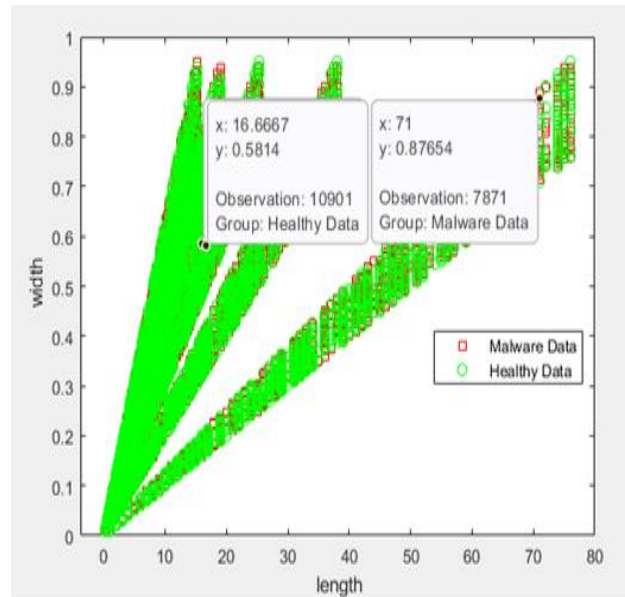


Figure 10:- Observation Of Healthy Data And Malware Data

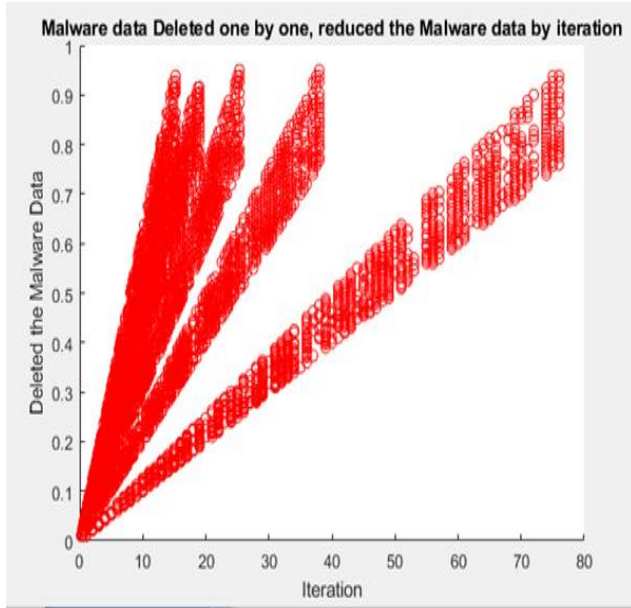


Figure 11:- malware data delation

It lets us make better use of our data and gives us a lot more information about how well our algorithm is working. Cross-validation is a resembling method that is used to test machine learning models with a small set of the available data. a single parameter called "k" that tells how many groups should be made from the data sample given. For the k-fold cross-validation, all that is needed is to use the cross-validation method more than once and then give the average result from all folds and runs. The standard error gives us an idea of how accurate this mean result will be as a measure of the model's true underlying mean performance on the dataset, which we don't know.

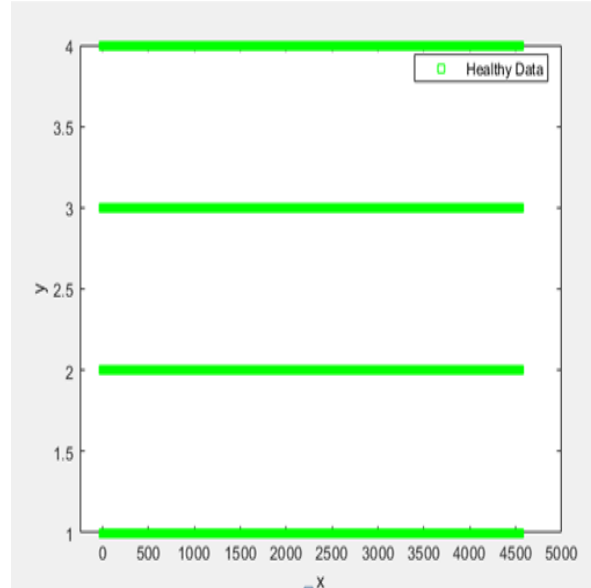


Figure 12:- seprate the malware and healthy data

Table 1 IoTQOS Result

Parameters values	Parameters values
Avg Delay	206.118855
Avg Energy	84.594327
Avg PDR	274.825140
Avg Through	206.118855

The Internet of Things (IoT) is a technology that is growing quickly and is going through a lot of changes right now so that people can get better services. In this survey report, the architecture explains what each layer of the Internet of Things does, and the taxonomy explains the factors that determine the quality of the services. Also, the metrics that need to be worked on to improve the quality of the services as a whole were given. In the future, this can be expanded by suggesting new ways to deal with problems like unstable links, traffic, node failure, packet loss, lifetime of node, congestion, and other similar problems in order to improve Quality of Service in the Internet of Things (IoT). Only a few different metrics were looked at in this work. In the future, the metrics can be explained in greater detail.

Table 2 Compare the Proposed Results with Existing Work

Parameters	Proposed Work	Existing Work[1]
	AODV-KNN	RL-QRP
Avg Delay	290.65	351
Avg Energy	173.05	196.3
Avg PDR	144.15	136.7
Avg Through	108.11	136.7

5. CONCLUSION

The proposed method increases security from source to destination without hurting the network's lifespan. However, there is still room for improvement in terms of cutting energy use along both the transmission and reception paths. The study of algorithms that are used to keep track of where the source and sink nodes are shows that quality of service, energy efficiency, and security are the factors that determine how well IoT applications can be used by end users and how likely they are to be adopted by them. To meet the Internet of Things (IoT) requirements for network security, both data security and protecting the locations of nodes that send and receive data are needed. Internet of Things devices that run on batteries have energy limits that have a big effect on how security algorithms are put into place. We've talked about some of the things that can affect how long IoT networks last and how safe they are. Concerns about the Internet of Things' quality of service (QoS) and security can be addressed with technologies like cognitive radio networks, fog computing, and machine learning. After figuring out what the end customers want, service providers can use these technologies to build real-time applications. Even though the level of security provided by different apps can be very different, any breach in security or attempt by bad actors to get through it can cost the end users a lot of money. When it comes to health care applications, a lack of security can put a person's life in danger. When it comes to military applications, a lack of security can be a threat to the security of the whole country. In the future, we will try to use techniques like machine learning to protect IoT networks and improve service quality.

REFERENCES

- [1] Manisha Bhatnagar Dolly Thankachan Identifying the Effects of Security Measures on QoS Variations for IoT Network: An Application Perspective Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021). IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4
- [2] F. A. Alaba, M. Othman, I. A. T. Hashem, and F. Alotaibi. "Internet of Things security: A survey". In: Journal of Network and Computer Applications 88 (2017), pp. 10–28.
- [3] C. Alcaraz, P. Najera, J. Lopez, and R. Roman. "Wireless sensor networks and the internet of things: Do we need a complete integration?" In: 1st International Workshop on the Security of the Internet of Things (SecIoT'10). 2010.
- [4] J. Arkko, D. Thaler, and D. McPherson. "IETF RC 7452: architectural considerations in smart object networking". In: IETF, Fremont, US (2015).
- [5] M. A. Bhabad and S. T. Bagade. "Internet of things: architecture, security issue and countermeasures". In: International Journal of Computer Applications 125.14 (2015).
- [6] G. Breed. "Wireless ad hoc networks: basic concepts". In: High frequency electronics 1 (2007), pp. 44–47.
- [7] L.-H. Chang, T.-H. Lee, S.-J. Chen, and C.-Y. Liao. "Energy-efficient oriented routing algorithm in wireless sensor networks". In: 2013 IEEE International Conference on Systems, Man, and Cybernetics. IEEE. 2013, pp. 3813–3818.
- [8] M. Chernyshev, Z. Baig, O. Bello, and S. Zeadally. "Internet of things (iot): Research, simulators, and testbeds". In: IEEE Internet of Things Journal 5.3 (2017), pp. 1637–1647.
- [9] A. Chhabra, V. Vashishth, A. Khanna, D. K. Sharma, and J. Singh. "An energy efficient routing protocol for wireless internet-of-things sensor networks". In: arXiv preprint arXiv:1808.01039 (2018)
- [10] Hamid G Alaoui, Z. El Belrhiti El Alaoui, A. FMG MAC: A fastG mobility adaptive MAC protocol for wireless sensor networks. Trans Emerging Tel Tech. 2020; 31:e3782. <https://doi.org/10.1002/ett.3782>
- [11] Kumar, S, Sharma, B, Singh, AK. An efficient algorithm for backbone construction in cognitiverradionetworks.IntJCommunSyst.2020;33:e4345. <https://doi.org/10.1002/dac.4345>

- [13] DadashiGavaber, Morteza, Rajabzadeh, Amir, "BADEP: Bandwidth and delay efficient application placement in fog-based IoT systems", Transactions on Emerging Telecommunications Technologies, 2020
- [14] Gomes, PH, Krishnamachari, B. TAMUĞ RPL: Thompson sampling based multichannel RPL. Trans Emerging Tel Tech. 2020; 31:e3806.
- [15] H. Liang, S. Yang, L. Li, and J. Gao. "Research on routing optimization of WSNs based on improved LEACH protocol". In: EURASIP Journal on Wireless Communications and Networking 2019.1 (2019), p. 194.
- [16] Dubey, K., Dubey, R., Panedy, S., & Kumar, S. (2024). A review of IoT security: Machine learning and deep learning perspective. *Procedia Computer Science*, 235, 335–346. <https://doi.org/10.1016/j.procs.2024.04.034>
- [17] Ghaffari, A., Jelodari, N., Pouralish, S., Derakhshanfard, N., & Arasteh, B. (2024). Securing Internet of Things using machine and deep learning methods: A survey. *Cluster Computing*, 27, 9065–9089. <https://doi.org/10.1007/s10586-024-04509-0>
- [18] Ni, C., & Li, S. C. (2024). Machine learning enabled Industrial IoT security: Challenges, trends and solutions. *Journal of Industrial Information Integration*, 39, 100549. <https://doi.org/10.1016/j.jii.2023.100549>
- [19] Wakili, A., Bakkali, S., & El Hilali Alaoui, A. (2024). Machine learning for QoS and security enhancement of RPL in IoT-enabled wireless sensors. *Sensors International*, 5, 100289. <https://doi.org/10.1016/j.sintl.2024.100289>
- [20] Sonavane, S. M., Prashantha, G. R., Nikam, P. D., R., M. A. V., & Bavirisetti, D. P. (2024). Optimizing QoS and security in agriculture IoT deployments: A bioinspired Q-learning model with customized shards. *Heliyon*, 10(2), e24224. <https://doi.org/10.1016/j.heliyon.2024.e24224>