DEEP LEARNING AND DATA MINING APPLICATIONS IN THE CYBERSECURITY PARADIGM TO FIGHT CYBER-ATTACKS

K.Thangadurai[1], S.Jayaprakash[2], M.Mailsamy[3]

[1,3]Associate Professor, Department of Computer Science and Engineering,AEC,Salem

[2] AssistantProfessor, Department of Computer Science and Engineering,AEC,Salem

Abstract:

This study examines a high-potential routing algorithm for cyber security that can be used in WSN-based IoT applications with huge traffic volumes. DM was the most effective and cutting-edge tool for discovering previously undetected important trends and patterns to boost an employee's performance. Data mining expertise is becoming more and more crucial for all firms. Data collection helps identify previously undiscovered and highly profitable data in massive amounts of data. Finding new patterns in a massive amount of data was the main goal of database information discovery. It combines various fields, including algorithms, AI, and statics. Users can view raw data from various IoT-based Applications thanks to DM, which organises a huge collection of data into a logical framework and extracts important information. As a result, the Internet of Things (IoT) is a network of actual physical objects or things that are linked together by computers, devices, and networks in order to gather and share data. The BS informs the Cluster heads to carry out tasks like help and trust to start the global calculation, and the agent at each CH then gives the CMs instructions to start the local calculation.

Keywords: Energy Consumption, deep learning, cyber security, data.

1. INTRODUCTION

Today, big data is increasingly related to deep learning, with the majority of solutions relying on deep learning approaches to uncover anomalies hidden within enormous data sets. As a result of recent advancements in communication technology, people and objects are becoming increasingly intertwined. Because of the Internet's accessibility, a range of devices that can connect and share information can be linked. IoT is an unique concept that allows users to link a range of sensors and smart devices from all around the world to collect real-time data.

Many academics are interested in using deep learning to identify Distributed Denial of Service threats. As a result, the research field was active in protecting the software that protected the network from issues. The goal of this study is to determine the best deep learning technique for detecting a Distributed Denial of Service assault.DM was a broad procedure that could be used to any sort of data; more recent studies on the issue can be

found in [2], in which specialists examined DM and deep learning approaches for analysing medical data. In a study of classification methods across data streams, the author explores classic classification methodologies.

For business decision-making, data mining methods should be combined with IoT. As a result, the primary purpose of this study is to provide a complete description of a computational system that has undergone extensive testing for Iot based Applications. Research studies in using DM to tackle IoT applications might leverage the architecture presented in this study as a roadmap.

2.Literature Review

"A future where physical items are smoothly incorporated into the information network, and where physical objects may become active players in business processes," write S. Haller et al. Services are available to interact with these'smart objects' through the Internet, query their state and any related information while keeping security and privacy concerns in mind."
It employs a memory data structure known as the DIU to hold closed item-sets. If a fresh transmission arrives, the CARM's algorithm checks and window sliding update the CIS's assistance. If CRAM detects any missing values in sensor data, instead of producing all potential association rules, it creates a method that is closely related to the current sensor readings. Based on these principles and selected closed item-sets that include item values, CRAM generates estimated values.The bandwidth of all monitoring radar data was used to create a PT using the canonical approach, and then the tree was rearranged in high bandwidth order.

[8] DSARM, a centralised technique, was offered as a way to locate the lost sensor's readings. It employs a mining technique called the rule of association to find radars that record similar information numerous times in a window slider, known as related radars, and then uses data related radars to measure information from a radar. It was difficult to apply a mining technique like Apriori directly to sensor data due to the nature of radar data. As a result, the experts developed the DSARM system, which applies the Apriori algorithm to data streams provided by sensor nodes.

[16] Umadevi et al. The cornerstone for behavioural analytics, which tries to avoid damage, is data mining. Deep learning provides a probabilistic and prognostic strategy in the long run. Patterns, regularities, and abnormalities are detected using deep learning and data mining techniques, enabling for the prevention of cybersecurity attacks.

3. Proposed method

In protocols that rely on collaboration, implicit trust was always there. It works with DM networks in IoT routing procedures. As IoT networks grow in size, they become increasingly vulnerable to attacks, necessitating the development of a robust protection system. [9]

Finding proper cryptography for wireless sensor networks is a serious difficulty.DLTSAD was a strong routing method that identified pathways for E-E transversal packets that used the least amount of total energy while simultaneously enhancing hostile node detection. We presented a cryptography-based security approach to deploy Elliptic Curve Cryptography in IoT. Improving the decryption and encryption components of the method, which currently give outstanding stability. [12]

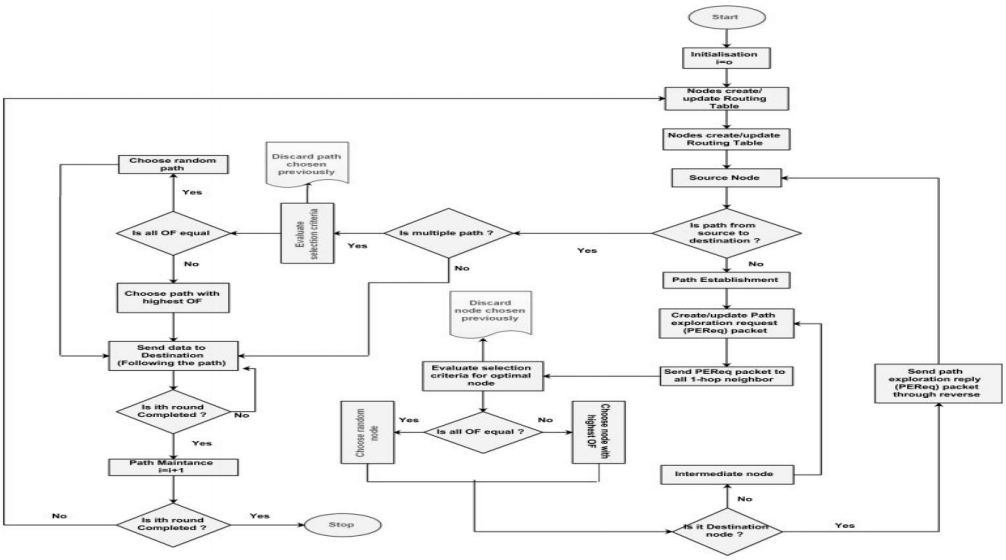


Figure 1:Sensor connection DM techniques

3.1. Proposed DLTSAD:

Algorithms analyse the dependability of connections while determining more secure routes. The most effective routes are determined by algorithms. [13]By selecting pathways that include units with high potential battery energy, an algorithm tries to lengthen the lifetime of a network. In addition, the algorithm calculates a higher number of security options.

3.1.1 DLTSAD Algorithm

This module provides DLTSAD, a threat monitoring and mitigation system that concentrates on collaboration, scheduling, and threat prevention and detection. We discuss the methodologies and approaches to trust-based protection in DLTSAD. This demonstrates how trust rely reasoning enables each joint to measure node performance and conduct a trust rely evaluation of the DLTSAD protocol using a unique language.

Algorithm 1: Proposed algorithm

Input : A network with N nodes, E links, Source node (N_o), Destination node (N_d)
Output : Multiple optimal paths from source to destination
Parameters:
OF: Optimality factor
L_E: Estimated lifetime of node
R_c: Reliability of communication
T_I: Traffic intensity of node
i: Round of algorithm
Initialize :
$i \leftarrow 0$
$R_c \leftarrow 0$
$T_I \leftarrow 0$
$L_E \leftarrow$ Estimated lifetime of node

```
begin
    while i≤100 && stopping criteria do
        Calculate OF of path;
        Create routing table of individual node;
        Source node checks path in its routing table;
        if path exists then
            Call algorithm 2 for sending data;
        else
            Call algorithm 3 for path discovery and establishment;
            Call algorithm 2 for sending data;
            Call algorithm 4 for path maintenance;
```

Algorithm 2: Algorithm for data forwarding

Input : Path
Output: Data successfully sent to destination

```
begin
    if path = multiple then
        Evaluate the selection criteria for optimal path and discard path chosen
          previously;
        if OF is equal then
            Choose random path
        else
            Choose path with highest
    Transmit data packet over the selected path;
```

Algorithm 3: Algorithm for path discovery and establishment

Input : Nodes
Output: Path from source to destination

```
begin
    while all 1-hop neighbor of source are not explored do
        Send the path exploration request (PEReq) packet to 1-hop neighbor;
        Check OF of its 1-hop neighbor node of node;
        if OF is equal then
            Choose random node;
        else
            Choose node with highest OF;
        Update the information about current node in the field of PEReq packet;
        if node = destination then
            Send path exploration reply (PERep) packet to source node;
        else
            Go to step 3;
```

Algorithm 4: Algorithm for path maintenance

Input : OF, L_E, R_c and , T_I
Output: Updated routing table

```
begin
    Calculate and update the values OF, L_E, R_c and , T_I in routing table of nodes
      with current value.
```

Figure 2 DLSAD algorithms

3.2 Data mining framework for cyber security:

For obtaining previously determined valuable discover patterns in order to enhance an employee's performance, DM was the most successful and rising technique. Data mining abilities are becoming increasingly important for all businesses. [15] In huge volumes of data, data mining helps in the uncovering of previously undiscovered and highly profitable information. A corporation could increase income by grouping things that are frequently purchased together, offering deals on some items, or eliminating duplicate items based on consumer purchasing habits.

.

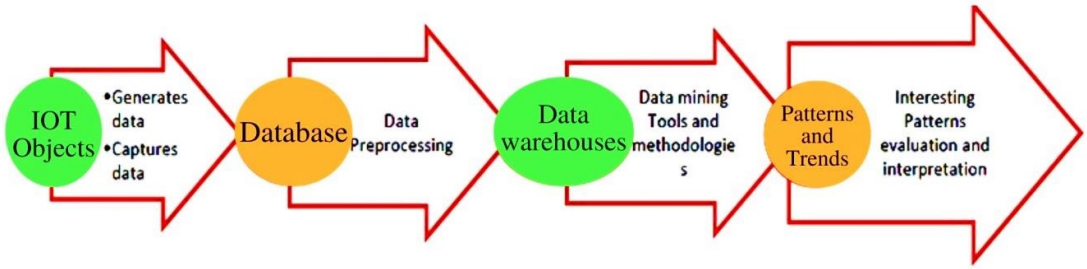


Figure 3: Data mining framework for IoT

3.2.1 Trust Computation:

We must first assess how much risk is suitable for each current operation before we can utilise the computed trust value to make a security decision. In other sense, a trust value criterion must be defined for each activity. The threshold node may be modified based on the security demands of each current operation. Comparing the anticipated reliability to the minimum trust model is a simple way to see if the trustee network fits the trust criterion.

3.2.2Phase of Support and Confidence Computing:

SCCP will be applied by all CHs who carry out the following procedures:

1. Keeping the enable item set and computing the candidate item set from the random item sets from the previous level at a later time.

2. Calculating the aid and trust rates.

3.4 Benefits

1. Extending the network's life cycle and ensuring a high degree of security.

2. It improves network performance while lowering overall energy usage. It also extends the network's lifespan.

3. The PDR and throughput ratio might both be improved.

4. Reduced average E-E latency and routing message overhead.

3.5 Architecture

These approaches use three parameters to simulate the best selection criteria. Durability, longevity node, and expected traffic intensity are the three factors.

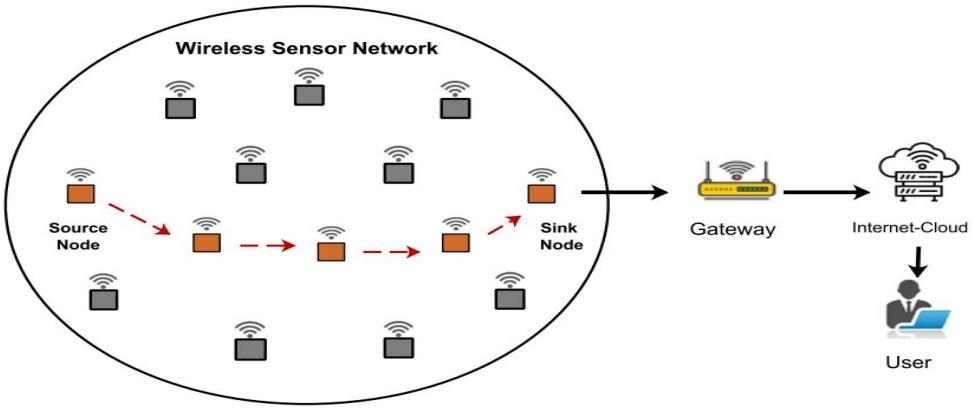


Fig. 4 WSN Architect

3.6 Model of Cyber Attack Procedure:

1. We must generate both a public and a private key as part of the key creation process. [22] The message should be decrypted by the receiver's private key and encoded by the sender's PK.

2. The PK was created using the formula below.

d * P = Q

3. The arbitrary number 'd'was picked from the range of 1 to n-1. P was the curve's starting point.

4. The private key was d, and the public key was Q. (public key).

4. RESULTS AND DISCUSSION

4.1 *Encryption / Decryption*:

The message should be displayed on the arc in encryption. Big data processing is included in the encrypted content. Examine the 'M' point on the 'E' curve for'm.' Pick 'k' from the table at arbitrary;

$$[1 – (n-1)] \quad (2)$$

C1 and C2 are the two cypher texts that will be created.

$$C1 = k*P \quad (3)$$

$$C2 = M + k*Q \quad (4)$$

C1 and C2 will be the ones to send.

Decryption refers to recovering the message m that was sent to the customer.

$$M = C2 – d * C1 \quad (5)$$

The first message, M, was sent to everyone.

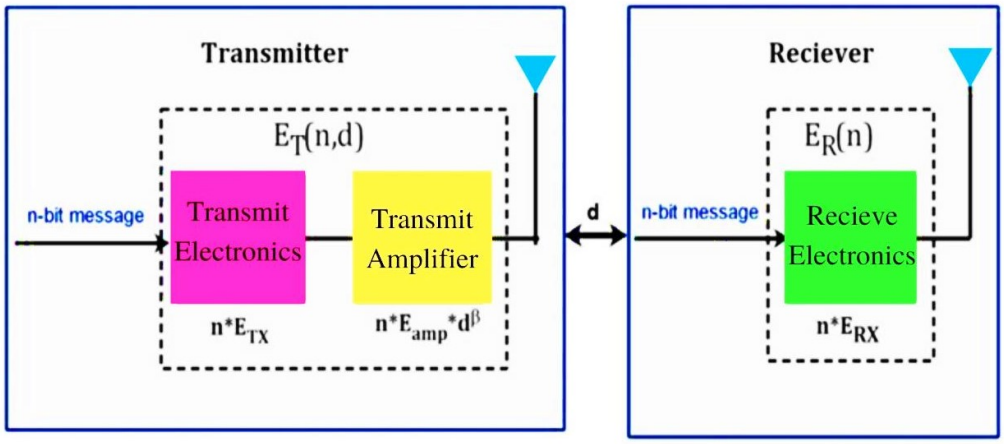


Fig. 6 Proposed protocols

4.2 *Throughput ratio*:

The throughput of a WSNs is defined as the number of properly transferred packets from sender to the receiver per second. A well-designed system should have a significant importance, and if it is targeted, the value of bandwidth will plummet.

Table 1 Simulation Results

Factors	Quantity
Simulation Time	10000ms
No. of Nodes	11
Packet	TCP
Protocol	AODV
Simulator	NS-2.4
Malicious node	2
Operating Platform	Ubuntu

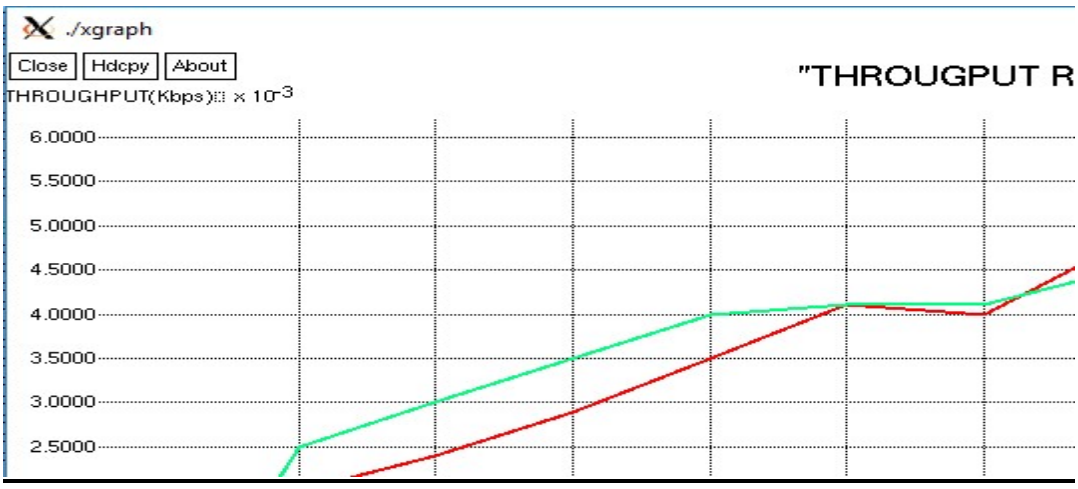


Fig. 7 Throughput ratio

4.3 *PDR:*

The ratio of total packets transmitted from an origin node to a target node in a network is referred to as PDR. The target should receive the highest packets of data possible. As the PDR value grows, so does the network output. PDR was calculated by comparing the network with and without a black hole hazard. The PDR was found to be extremely poor during the attack compared to before the attack, meaning that fewer packets were sent to the mobile sink.

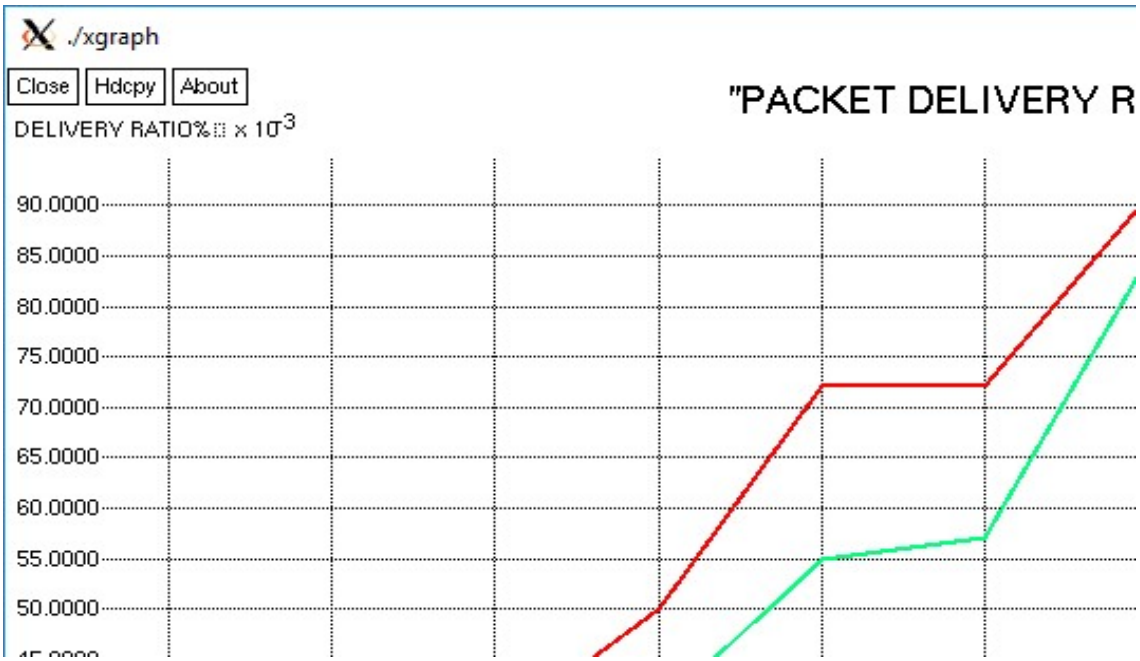


Figure 8 PDR

4.4 *Energy Consumption:*

Energy evaluation is critical for determining the requirements of an intense data process that runs smoothly on mobile devices. This research presents an experimental investigation of the energy usage of DM algorithms running on mobile devices.

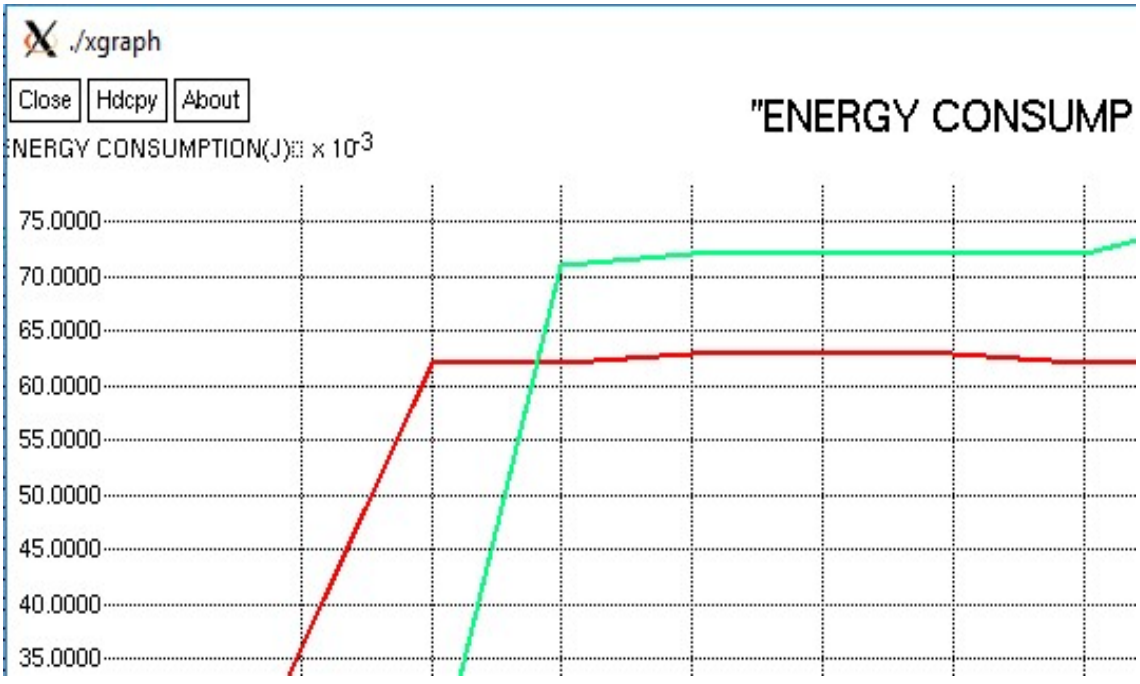


Figure 9 Energy Consumption

4.5 *E-E delay:*

The packet's end-to-end latency is the sum of the delays observed at a succession of transitional nodes on the route to the destination. A fixed broadcast and dispersion delay and a changeable processor and queuing time at the nodes make up each delay.

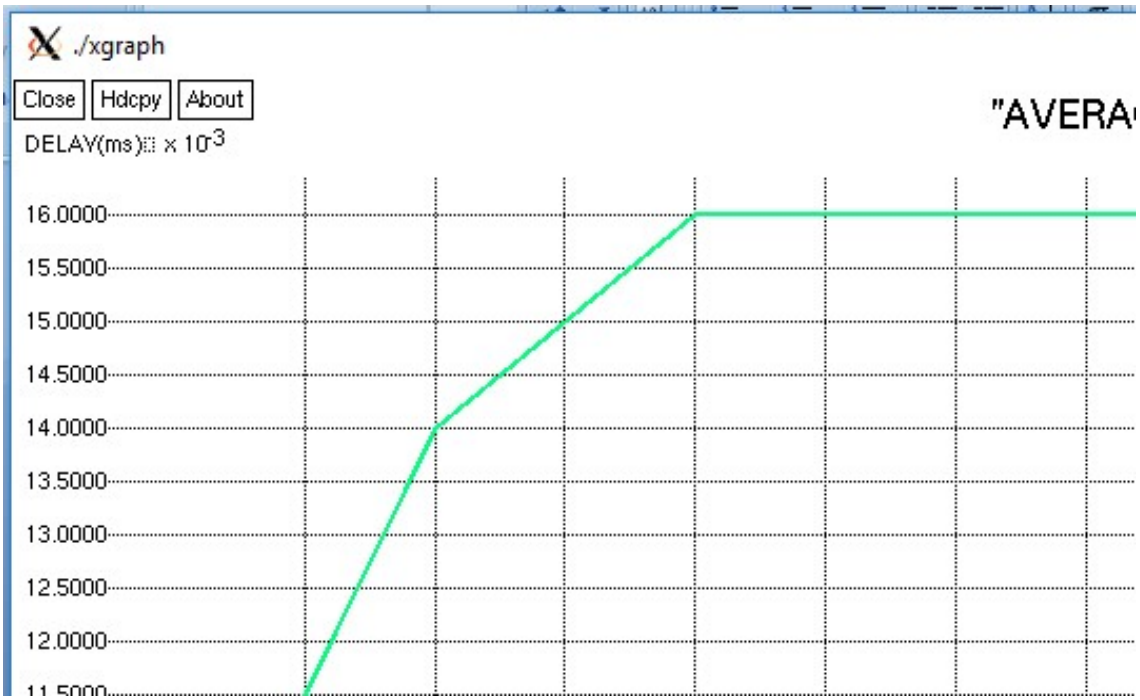
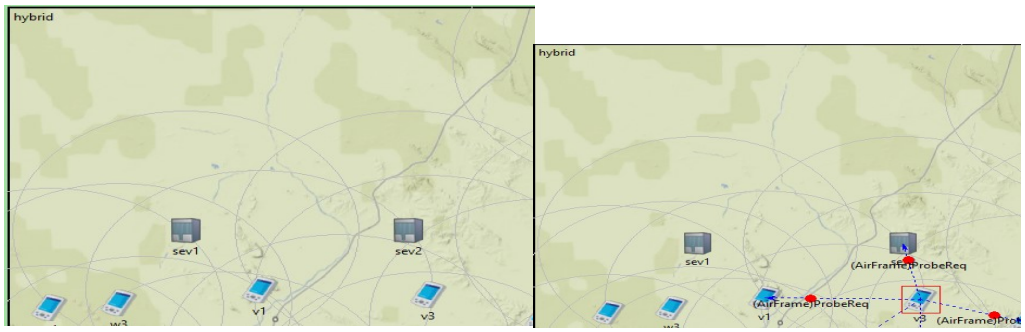


Fig. 10 E-E delay

V. OUTPUT



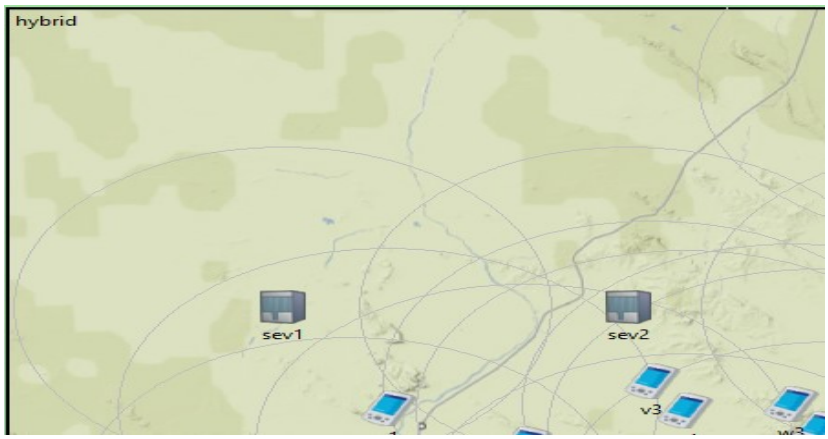


Fig. 11 outputs

4. CONCLUSION

A flurry of algorithms has sprung up in response to the increased need for DM approaches in the realm of WSNs. These methods address issues relating to the design and execution of WSNs. This research looks at a high-potential virtual routing algorithm that can be employed in WSN-based Iot based Applications with significant traffic volumes. In order to enhance an organisation 's performance, DM was the most efficient and emerging technology for extracting completely undiscovered valuable trends and patterns. Data mining skills are becoming increasingly important to all businesses.

Reference:

1. Sarker, Iqbal H., et al. "Cybersecurity data science: an overview from deep learning perspective." *Journal of Big data* 7.1 (2020): 1-29.
2. Bhamare, Deval, et al. "Cybersecurity for industrial control systems: A survey." *computers & security* 89 (2020): 101677.
3. KABANDA, GABRIEL. "Performance of Deep learning and other Artificial Intelligence Paradigms InCybersecurity." *Oriental journal of computer science and technology* 13.1 (2020): 1-21.
4. Rekha, Gillala, et al. "Intrusion detection in cyber security: role of deep learning and data mining in cyber security." *Advances in Science, Technology and Engineering Systems Journal* 5.3 (2020): 72-81.
5. Khan, Shah Khalid, et al. "Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions." *Accident Analysis & Prevention* 148 (2020): 105837.

6. Kabanda, Gabriel. "A bayesian network model for deep learning and cyber security." *Proceedings of the 2nd Africa-Asia Dialogue Network (AADN) International Conference on Advances in Business Management and Electronic Commerce Research*. 2020.
7. Haider, Noman, Muhammad ZeeshanBaig, and Muhammad Imran. "Artificial Intelligence and Deep learning in 5G Network Security: Opportunities, advantages, and future research trends." *arXiv preprint arXiv:2007.04490* (2020).
8. Samtani, Sagar, et al. "Cybersecurity as an industry: A cyber threat intelligence perspective." *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (2020): 135-154.
9. Ayodeji, Abiodun, et al. "A new perspective towards the development of robust data-driven intrusion detection for industrial control systems." *Nuclear engineering and technology* 52.12 (2020): 2687-2698.
10. Alsaedi, Abdullah, et al. "TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems." *IEEE Access* 8 (2020): 165130-165150.
11. Alghamdi, Mohammed I. "Survey on Applications of Deep Learning and Deep learning Techniques for Cyber Security." *International Journal of Interactive Mobile Technologies* 14.16 (2020).
12. Gupta, Maanak, et al. "Security and privacy in smart farming: Challenges and opportunities." *IEEE Access* 8 (2020): 34564-34584.
13. Rath, Mamata, and Sushruta Mishra. "Security approaches in deep learning for satellite communication." *Deep learning and data mining in aerospace technology*. Springer, Cham, 2020. 189-204.
14. Coulter, Rory, et al. "Code analysis for intelligent cyber systems: A data-driven approach." *Information sciences* 524 (2020): 46-58.
15. Truong, Thanh Cong, et al. "Artificial intelligence and cybersecurity: Past, presence, and future." *Artificial intelligence and evolutionary computations in engineering systems*. Springer, Singapore, 2020. 351-363.
16. Gunduz, MuhammedZekeriya, and Resul Das. "Cyber-security on smart grid: Threats and potential solutions." *Computer networks* 169 (2020): 107094.
17. Coulter, Rory, et al. "Data-driven cyber security in perspective—Intelligent traffic analysis." *IEEE transactions on cybernetics* 50.7 (2019): 3081-3093.
18. Rawat, Danda B. "Journal of Cybersecurity and Privacy: A New Open Access Journal." *Journal of Cybersecurity and Privacy* 1.1 (2021): 195-198.
19. Zhang, Jun, et al. "Deep learning based attack detection for cyber-physical system cybersecurity: A survey." *IEEE/CAA Journal of AutomaticaSinica* 9.3 (2021): 377-391.
20. Elsisi, Mahmoud, et al. "Towards secured online monitoring for digitalized GIS against cyber-attacks based on IoT and deep learning." *Ieee Access* 9 (2021): 78415-78427.