**Webology**, Volume 18, Number 4, 2021
**ISSN:** 1735-188X
**DOI:** 10.29121/WEB/V18I4/128

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

**Surendra Shukla[1], Bhasker Pant[2], Dibyahash Bordoloi[3]**

[1]Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India
[2]Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India
[3]Head of the Department, Department of Computer Science & Engineering,Graphic Era Hill University, Dehradun, Uttarakhand India

## ABSTRACT

In this study, a high-potential routing algorithm for cyber security that may be applied to IoT applications based on WSNs and with high traffic volumes is investigated. The most effective and cutting-edge method for identifying crucial, previously unnoticed trends and patterns to improve an employee's performance was dynamic modelling (DM). Knowledge of data mining is becoming increasingly important for all businesses. 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. 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.

## 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.

## Literature Review

Services are available to query the state of these "smart objects" and any relevant information through the Internet while keeping security and privacy considerations 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. 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.

[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.

## 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]

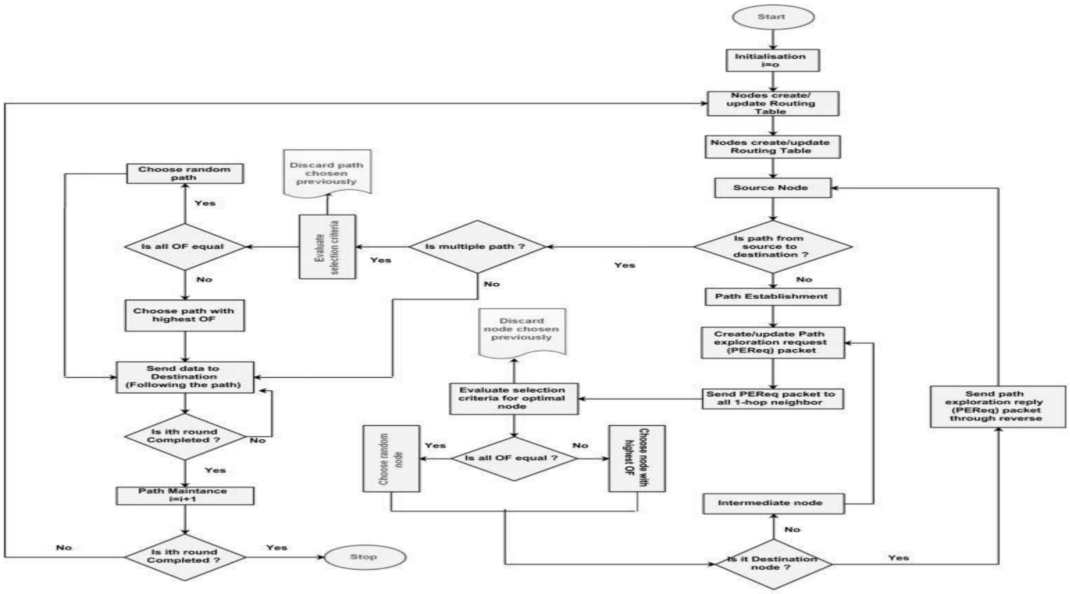


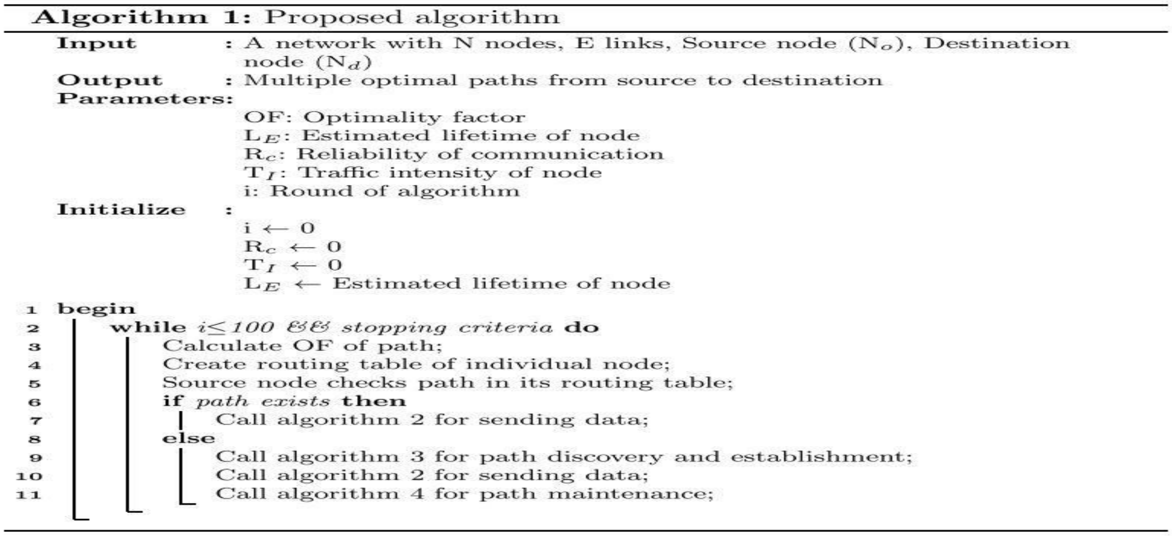
*Figure 1: Sensor connection DM techniques*

### 3.1. Proposed DLTSAD:

As they choose safer routes, algorithms analyse the dependability of links. Algorithms are used to select the best routes. [13] The programme also calculates a greater variety of security alternatives.

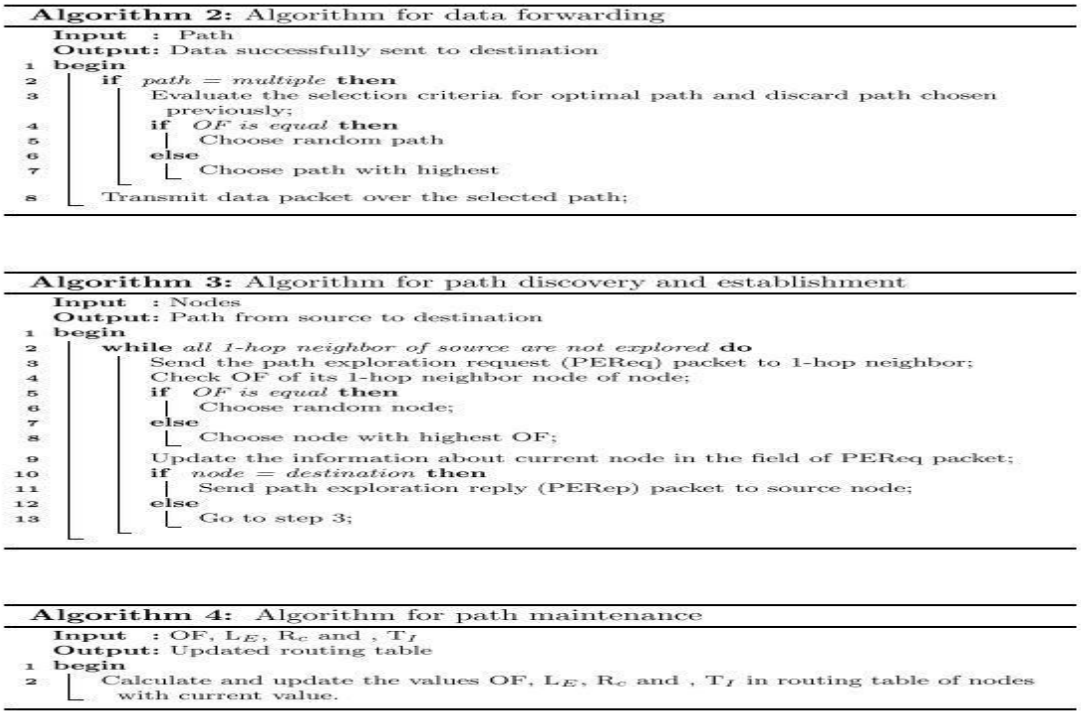
### 3.1.1 DLTSAD Algorithm

This module offers DLTSAD, a system for threat monitoring and mitigation that focuses on coordination, planning, and the prevention and detection of threats. We go over the methods and strategies for trust-based security in DLTSAD. This exemplifies how each joint can measure node performance and carry out a trust rely evaluation of the DLTSAD protocol using a special language thanks to trust rely reasoning.



```
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 $i \leq 100$ && 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;
```

**Webology**, Volume 18, Number 4, 2021
**ISSN:** 1735-188X
**DOI:** 10.29121/WEB/V18I4/128



```
Algorithm 2: Algorithm for data forwarding
   Input  : Path
   Output: Data successfully sent to destination
1  begin
2      if path = multiple then
3          Evaluate the selection criteria for optimal path and discard path chosen
             previously;
4          if OF is equal then
5              Choose random path
6          else
7              Choose path with highest
8      Transmit data packet over the selected path;
```

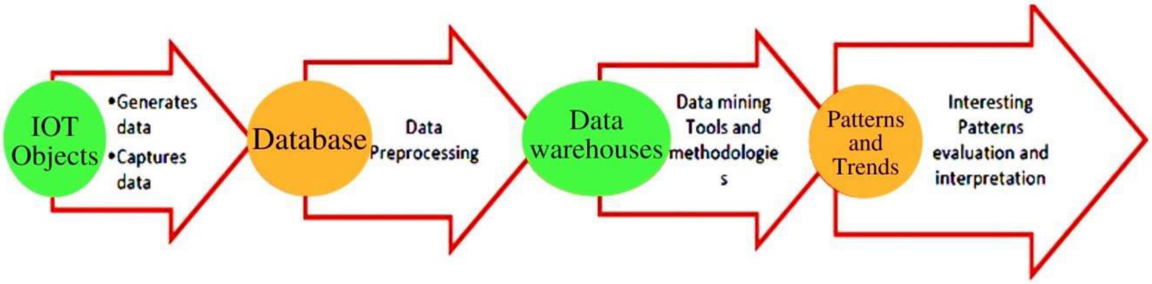
```
Algorithm 3: Algorithm for path discovery and establishment
   Input  : Nodes
   Output: Path from source to destination
1  begin
2      while all 1-hop neighbor of source are not explored do
3          Send the path exploration request (PEReq) packet to 1-hop neighbor;
4          Check OF of its 1-hop neighbor node of node;
5          if OF is equal then
6              Choose random node;
7          else
8              Choose node with highest OF;
9          Update the information about current node in the field of PEReq packet;
10         if node = destination then
11             Send path exploration reply (PERep) packet to source node;
12         else
13             Go to step 3;
```

```
Algorithm 4: Algorithm for path maintenance
   Input  : OF, $L_E$, $R_c$ and , $T_I$
   Output: Updated routing table
1  begin
2      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] Data mining assists in locating previously unknown and very valuable information in massive data sets. Grouping often purchased items together, providing discounts on particular items, or getting rid of identical products depending on client preferences could all help a business make more money.



.

*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.

### 3.2.2 Phase of Support and Confidence Computing:

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

1. Preserving the enable item set and determining the candidate item set later by using the random item sets from the previous level.
2. Figuring out the trust and assistance rates.

### 3.4 Benefits

1. Increasing security and extending the network's life cycle.
2. It reduces total energy consumption while enhancing network performance. The lifespan of the network is also increased.
3. There may be room to raise both the PDR and throughput ratio.
4. Reduced overhead in message routing and average E-E delay.

### 3.5 Architecture

To simulate the best selection criterion, these approaches require three factors. The three factors are durability, lifetime of node, and anticipated traffic volume.

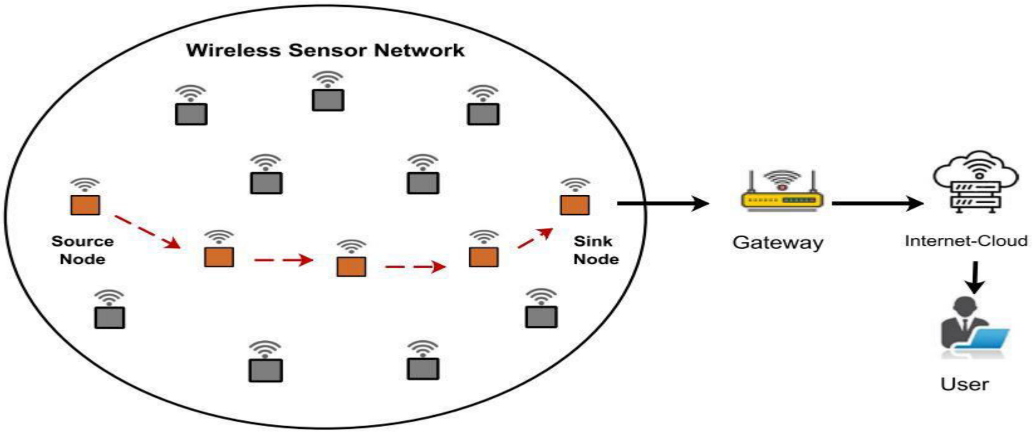


*Fig. 4 WSN Architect*

### 3.6 Model of Cyber Attack Procedure:

1. As part of the key creation procedure, we need to generate both a public and a private key. [22] The recipient's private key should be used to decode the message, and the sender's PK should be used to encode it.
2. The following formula was used to construct the PK.

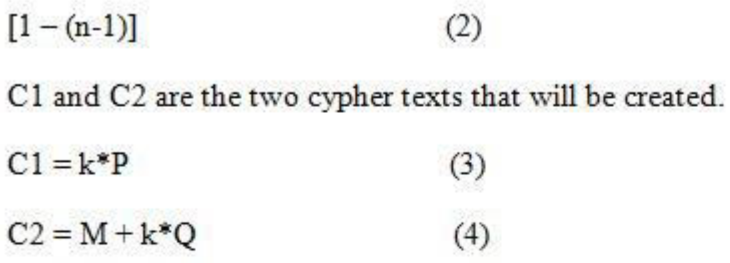
d * P = Q 3. 'd' was chosen at random from a range of 1 to n-1. P was the beginning of the curve.

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

## Results And Discussion

### 4.1 *Decryption or Encryption*:

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;



$$[1-(n\text{-}1)] \qquad (2)$$

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

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

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

C1 and C2 will be the ones to send.

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

$$M = C2 - d * C1 \qquad (5)$$

The first message, M, was sent to everyone.

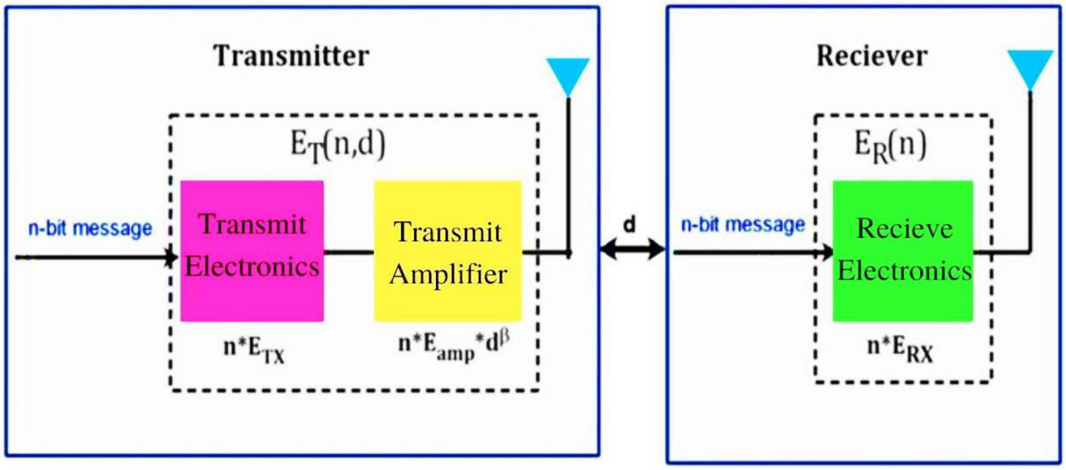


*Fig. 6 Protocols*

### 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. If a well-designed system is intended, the value of bandwidth will decline because it should be of substantial importance.

*Table 1 Results*

| Attributes | Quantity |
|---|---|
| Simulation Time | 10000ms |
| No. of Nodes | 11 |
| Packet | TCP |
| Protocol | AODV |
| Simulator | NS-2.4 |

| | |
|---|---|
| **Malicious node** | 2 |
| **Operating Platform** | Ubuntu |

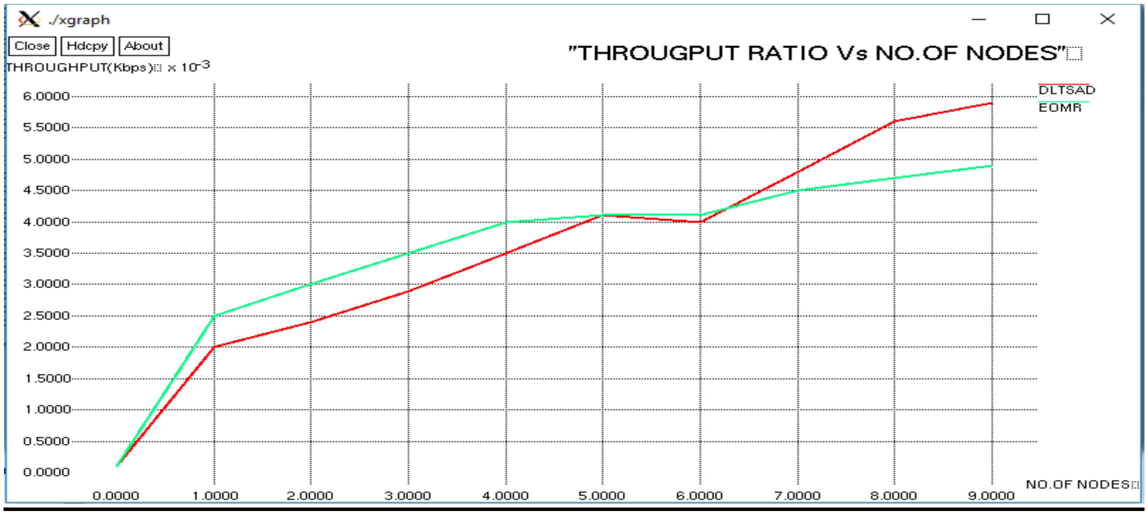


*Fig. 7 Throughput ratio*

### 4.3 *PDR:*

PDR in a network is the proportion of all packets sent from an origin node to a target node. The target should get as many data packets as is practical. The network output increases in tandem with the PDR value. By contrasting the network with and without a black hole hazard, PDR was determined. Fewer packets were sent to the mobile sink as a result of the attack's extremely poor PDR as compared to before the attack.

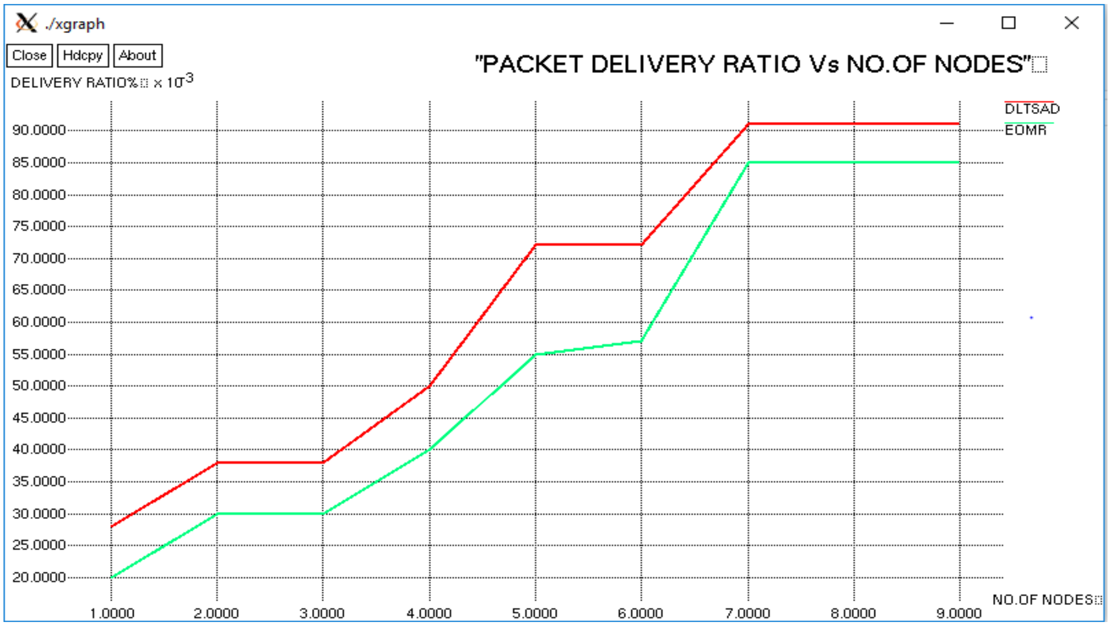


*Figure 8 PDR*

### 4.4 *Consumption of Energy:*

Analyzing energy is essential for figuring out what is needed for a demanding data process to function properly on mobile devices. In this study, the energy consumption of mobile device-running DM algorithms is experimentally investigated.

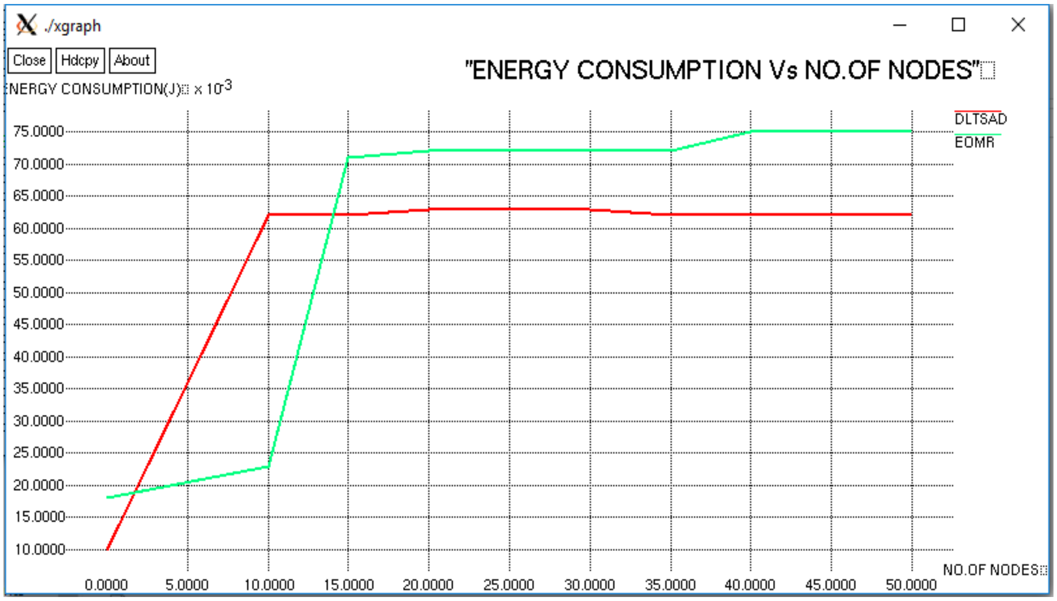


*Figure 9 Consumption of Energy*

### 4.5 *E-E delay:*

The packet's end-to-end latency is calculated as the total of the delays experienced at each transitional node on the route to the destination. The broadcast and dispersion delays are fixed, and the processing and queuing times at the nodes are variable.

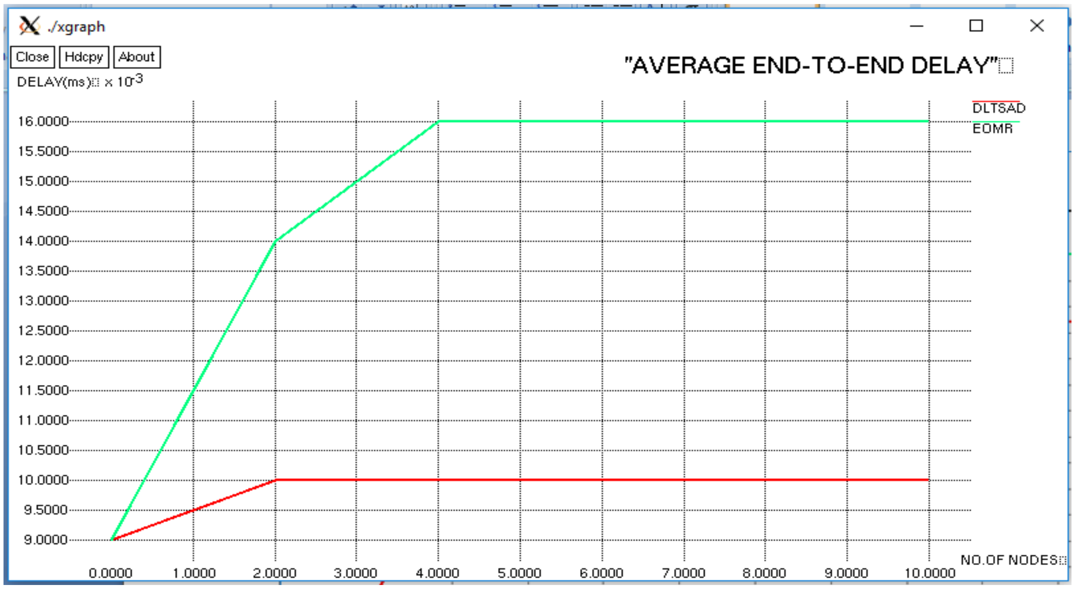


*Fig. 10 E-E delay*

## Result

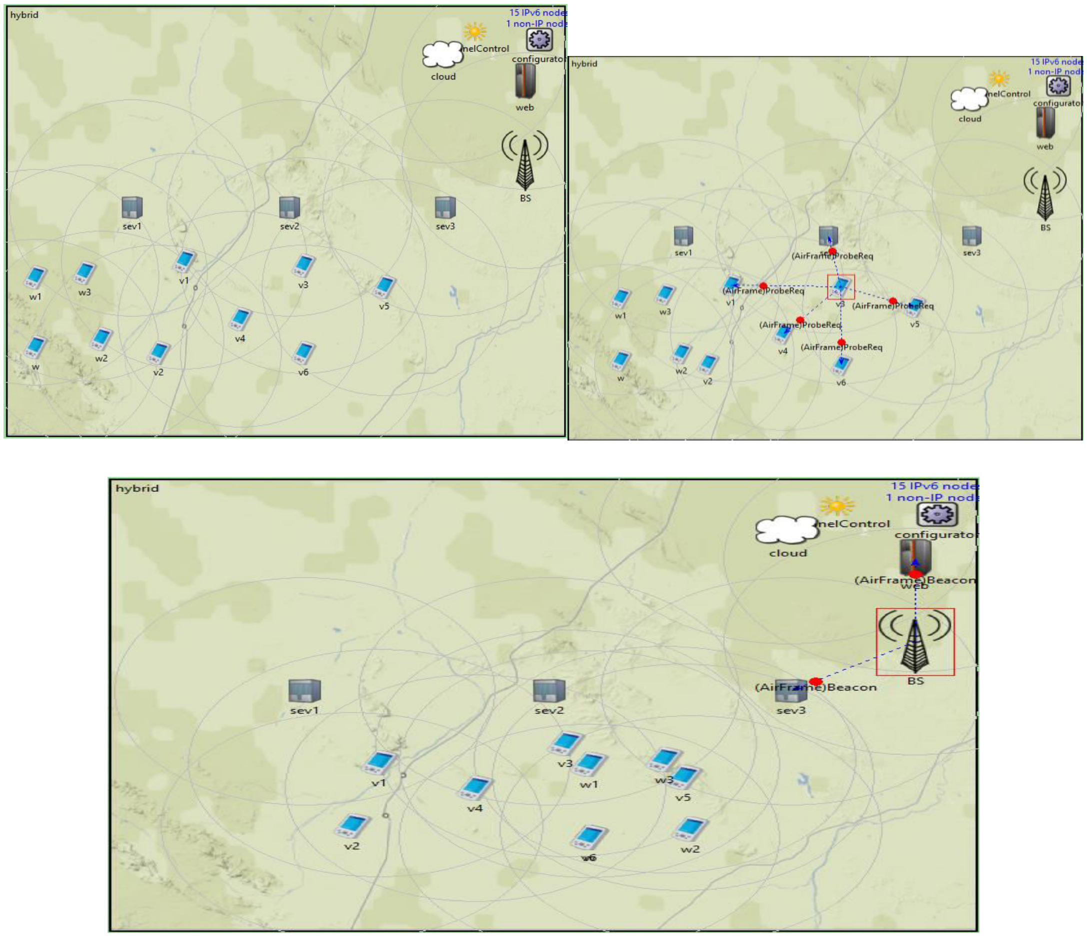


*Fig. 11 results*

## 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. 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 In Cybersecurity." *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 Zeeshan Baig, 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, Muhammed Zekeriya, 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 Automatica Sinica* 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.

21. Haji, Saad Hikmat, and Siddeeq Y. Ameen. "Attack and anomaly detection in iot networks using deep learning techniques: A review." *Asian journal of research in computer science* 9.2 (2021): 30-46.
22. Larriva-Novo, Xavier, et al. "Efficient distributed preprocessing model for deep learning-based anomaly detection over large-scale cybersecurity datasets." *Applied Sciences* 10.10 (2020): 3430.