International Journal of Education and Science Research Review

E-ISSN 2348-6457

July –August 2017, Volume-4, Issue-4

P-ISSN 2349-1817

www.ijesrr.org Email- editor@ijesrr.org

DYNAMIC SINK RELOCATION STRATEGIES FOR OPTIMIZED DATA COLLECTION IN LARGE-SCALE IOT NETWORKS

Udathu NVP Rajendranath, Research Scholar, Sunrise University, Alwar

Dr. R. Murali Prasad, Professor, Sunrise University, Alwar

ABSTRACT

Due to sensor energy consumption being non-uniform and potentially reducing network lifespan, data transmission from cluster head to sink node across many hops is undesirable. This led to the development of the idea of sink mobility. Dynamic Sink Mobility for Data gathering (DSMDC) is a suggested method for fully using sink mobility, which combines effective data gathering with random moves. Detected Event Frequency (DEF) is the foundation for sink migration.

Keywords: Internet, transmission, gadgets, network, Dynamic.

INTRODUCTION

In the last ten years, the proliferation of interconnected devices—particularly those that make use of the Internet of Things (IoT)—has caused a meteoric rise in the volume of data created all over the world. With an exponential growth rate anticipated, Cisco predicted that the Internet of Things (IoT) would produce over 500 zettabytes of structured and unstructured data per year by the end of 2019. In addition, estimates from a variety of sectors indicate that there will be 50 billion Internet-enabled gadgets by 2020.

In addition, the proliferation of smart information-aware devices in recent years—including sensors, actuators, smartphones, smart wristbands, tablets, RFID reader devices, and Machine-to-Machine (M2M) connections—has contributed to the exponential growth of the volume of data generated. Industry and academia alike use these devices—and sensors in particular—to gather information for a wide range of purposes, including healthcare, environmental monitoring, and precision agriculture [10]. Innovative data gathering, transmission, storage, and processing architectures have been established as a result of substantial advancements in engineering and research. In addition, there has been a resurgence of interest in the Big Data paradigm in both academia and business [11]. This paradigm is mostly used to explain enormous data sets. Data science, the Internet of Things (IoT), and artificial intelligence are just a few areas that have reaped the rewards of Big Data's digital revolution.

In addition, tech giants like Amazon, Google, Facebook, and Twitter are driving the present wave of Big Data studies. These studies employ large-scale data mining and analysis to create a slew of value-added services using real-time data sourced from humans, such as e-mails, online purchase histories, and tweets [12]. Big Data, in its most basic definition, is a massive amount of data, which can be organized in a variety of ways and contains structured, semistructured, and unstructured information. This data is created at a high velocity and has the ability to shed light on hidden information within datasets, as well as assist with their management and organization [13].

Generating data, acquiring it, storing it, and analyzing it are the four main components of the Big Data process. Due to the inclusion of large volumes of unstructured data of many kinds, Big Data necessitates more real-time analysis than conventional data.
A wide variety of important application fields, such as eHealth, smart environments, smart cities, smart buildings, and precision agriculture, are now making frequent use of IoT devices. A plethora of sensor-

based applications have emerged as a result of the enormous data gathering made possible by IoT technology. Wireless sensor networks (WSNs) are the most popular IoT-based platforms for data collecting [14], because to the virtualization of smart products made possible by IoT technology. Therefore, WSNs are well-known to have been an essential part of what eventually became the Internet of Things (IoT) infrastructure. Thousands upon thousands of sensors work together in these networks to gather data and send it back to the central station.

LITERATURE REVIEW

Khan, A. W., Abdullah, A. H., Anisi, M. H., & Bangash, J. I. (2014). Many systems have recently taken use of sink mobility to extend the lifespan of WSNs. By eliminating energy-holes and promoting balanced energy consumption across nodes, mobile sink-based techniques improve upon standard WSNs that send sensory data from the sensor field to a static sink. Nodes in mobility situations must monitor the most recent positions of mobile sinks in order to ensure data transmission. The energy conservation aim is undermined by the frequent transmission of sink topological changes; therefore, management of this process is necessary. Controlled propagation of topological changes to sinks also impacts routing strategy performance, which in turn increases data delivery delay and decreases packet delivery ratios. Using sink mobility as an example, this research offers a taxonomy of data gathering and dissemination systems.

Vijayalaxmi et al. (2015) for the purpose of gathering data from the WSN via a mobile sink system. The nodes were grouped into clusters based on the K-medoid technique, which was used for clustering. The data came from a variety of sensor nodes, and the portable sink toured each cluster.

To define the mobile sink's route, Amar Kaswan et al. (2017) offered two algorithms: RkM (reduced k-means) and DBRkM (delay bound reduced k-means). The results of the studies demonstrated that both of the aforementioned algorithms efficiently chose a route for the mobile sink. Factors that impacted the algorithms were lowering the required RPs (rendezvous points) distance, decreasing the average journey duration, and improving the range of one-hop members. By linking sensor nodes via one-hop communication, RkM was able to ascertain the route. In the background, DBRkM developed a time-limited route using the same procedure. The algorithms presuppose that the mobile sink has a small time of stay and that each node has an equal load when it comes to creating information. In comparison to WRP and CB, the suggested algorithms demonstrated efficient results when measuring hop count, energy consumption, active node count, and network longevity. In addition, the authors suggested a system for gathering data from portable sinks. To keep packet loss to a minimum, the strategy may be used throughout each data gathering cycle.

Dynamic Sink Mobility for Data Collection

Model for the System

In this scenario, we think of a massive sensor network in which every node acts as either a source or a router. The source node is responsible for collecting environmental data. The data that has been seen should be sent to the sink node by the router node. The GPS-enabled assumption is that every sensor node knows its precise position. The sensor nodes engage in one-hop communication with the cluster head, and the cluster head in turn engages in one-hop communication with the sink. The suggested protocol's foundation is a virtual grid with a fixed number of nodes per cell. In Section B, the process of selecting cluster heads is detailed. With the use of single-hop communication, the cluster head compiles the data acquired by each sensor node and sends it to the sink. Sink is able to relocate to other grids. The suggested system's architecture is shown in Figure 1.

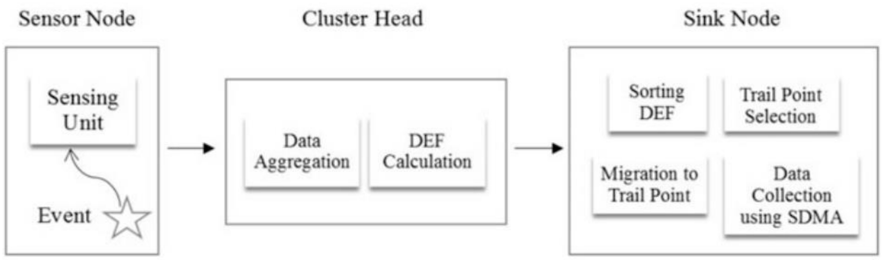

Fig. 1 System architecture

Head Selection for Clusters

A cluster head is chosen for each virtual grid once the network has been partitioned. At the outset, the network's nodes are all waiting for the cluster head candidacy notification to arrive. When it receives a message from a node in its own cluster, it appoints that node as the cluster head. Based on the remaining energy of the presently chosen cluster head, the function of cluster head is rotated across sensors. You may dynamically control the power of sensor nodes by setting two threshold values: Max_threshold and Min_threshold. The sensor node's Max_threshold represents its maximum accessible energy and Min_threshold its minimum remaining energy. When the current node's energy supply is low, you may choose another one by advertising Min_threshold. Cluster head selection is repeated whenever a node's energy surpasses Min_threshold.

Determining the Frequency of Detection Events

The number of events identified per cluster throughout the Mobility Time Period (MTP) is known as the detection event frequency. The cluster ID and the corresponding detection event frequency are stored in a table by the sink node. The data is sorted, and the set of grids with the most data is identified. After that, the sink relocates in accordance with the suggested methodology for sink migration.

Migrating Sinks

At the very heart of any gird construction is the sink node. It collects cluster information on discovered data for each MTP. Using the industry-standard SDMA protocol, the sink node simultaneously retrieves data from the cluster head. Here we present and evaluate three distinct sink migration techniques: DSMDC, DEF-A, and DEF-D. These algorithms are designed for data collection and use the detected event frequency in an ascending or descending order, respectively. In comparison to the other suggested techniques, DSMDC had a greater throughput while using less energy, and its network lifespan was also enhanced.

Algorithm 1: DSMDC

```
Input: Max_DEF
Output: Sink Migration Path
n ← Maximum number of grid set
D = { D1, D2, … Dg }     /* DEF of each grid set g */
Trail Points = { TP1, TP2 … TPg }
for each g in grid set
if Dg == Max_DEF
        TP1 ← g
        i←g
j←2
while j <= n
            TPj ← (i+1) mod n
            j ←j+1
            i ←i+1
return Trail Points;
```

Algorithm 2 Trail Point Finder

```
Input: DEF_grid (Cluster co-ordinates of a grid)
Output: Current_TrailPoint
IF x.DEF_grid <= 1
IF y.DEF_grid <=1
        Current_TrailPoint = TP_1
ELSE
        Current_TrailPoint = TP_2
ELSE
IF y.DEF_grid <=1
        Current_TrailPoint = TP_3
ELSE
            Current_TrailPoint = TP_4
END IF
return Current_TrailPoint
```

Algorithm 3 DEF-A

```
Input: D = { D1, D2,...Dn}     /* DEF of each grid*/
Output: Trail Points
n = number of grids
A = sorted array of D in Ascending order
i = 0
while i < n
TP = Trail Point Finder (DEF_grid)
i ← i + 1
```

Mobile Sink Scheme for Efficient Data Collection in the Internet of Things

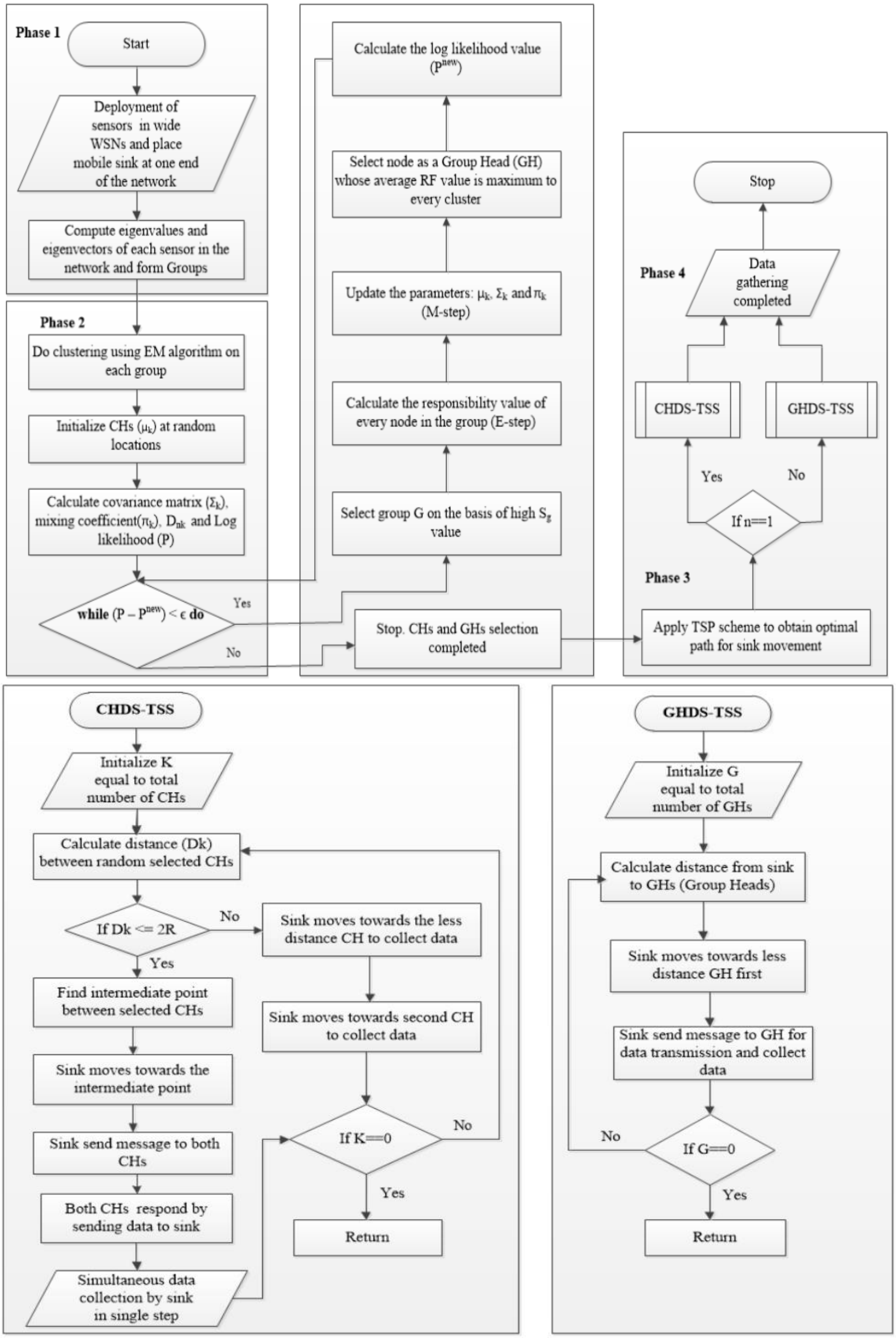

Fig. 2 Flowchart of the proposed methodology

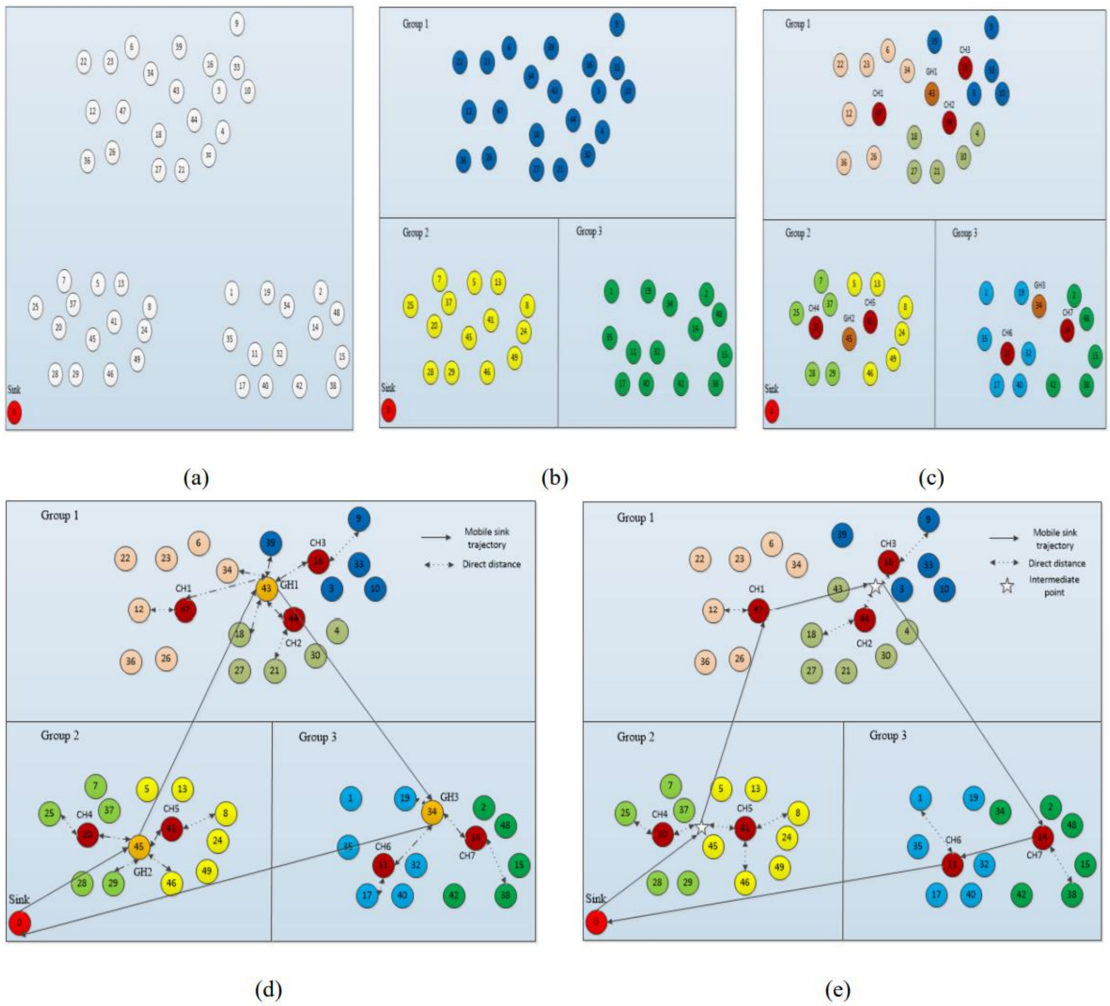

Fig. 3. (a) Deployment of sensor nodes (b) Groups formation (c) Clusters formation (d) Data collection using CHDS-TSS (e) Data collection using GHDS-TSS.

Performance Evaluation

The suggested model is tested using the NS2 (Network Simulator version 2) simulator to determine its overall performance. We test the algorithm's performance on many networks by simulating them with varying numbers of nodes on a 1000*1000 square meter (m2) surface. The starting energy of each sensor node is 2J, and there is no energy limitation for the mobile sink node. The specified speed for the mobile sink is 2 m/s. Table 1 provides a summary of all the simulation parameters for the suggested model.

Table-1: Parameters summary

Parameter	Value
Network area size	$1000\times1000m^2$
Range of sensor nodes	25, 40, …, 200
Size of data packet	512 bytes
Size of control packet	32 bytes
Initial energy of nodes	2J
Sink speed	2m/s
E_{elec}	50nJ/bit
ε_{fs}	$10pJ/bit/m^2$
ε_{mp}	$0.0013pJ/bit/m^4$
D	75m

In this comparison, the general EEM mobility based data collection strategy is pitted against the CHDS-TSS and GHDS-TSS suggested methods, which include data transmission from cluster nodes to cluster heads and subsequent sink movement towards each cluster head for data collection. When compared to the general EEM mobility strategy, the suggested methods vary primarily in their focus on reducing the overall duration of mobile sink excursions.

A. Metrics for Performance Analysis

Using commonly used performance indicators like packet delivery ratio, average energy consumption, latency, and network lifespan, we compare the conventional mobility method to data gathering with the suggested strategies.

PDR, or packet delivery ratio: Here we see the ratio of packets delivered and received by the sender and receiver, as a statistic. In essence, it reveals the percentage of packets successfully received at their final destination.

$$PDR = \frac{P_{received}\times 100}{\sum_{k=1}^{n} P_{generated\,k}}$$

Data transmission, reception, and processing all need energy, and this metric shows how much power each sensor node in the network model uses on average.

The average amount of time it takes for a packet to go from being created to being received at the sink is defined by the network performance indicator known as delay. The average latency of n sensor nodes is denoted by D(n).

$$D(n) = \frac{\sum_{k=1}^{P_{received\,k}} (T_{received\,k} - T_{transmitted\,k})}{P_{received}}$$

Lifespan of the network: It indicates how long the network will be operational at full capacity. When estimating a network's lifespan, it's possible to use the following time spans: when all of the sensors die at once, when a single sensor dies, or when the network is partitioned.

B: Discussion and Findings

This section describes the simulation findings that were produced by using the suggested methodologies and compares them to the basic EEM mobility data gathering method.

Waiting around

The simulation results show that compared to a general EEM technique, CHDS-TSS and GHDS-TSS provide better average delays. This is due to the fact that in a standard EEM setup, the sink will go to each cluster head in order to gather data. Therefore, it experiences a maximum delay in expansive and big network areas due to the slower sink movement speed compared to communication among sensor nodes, which causes it to take longer to reach a specific cluster head.

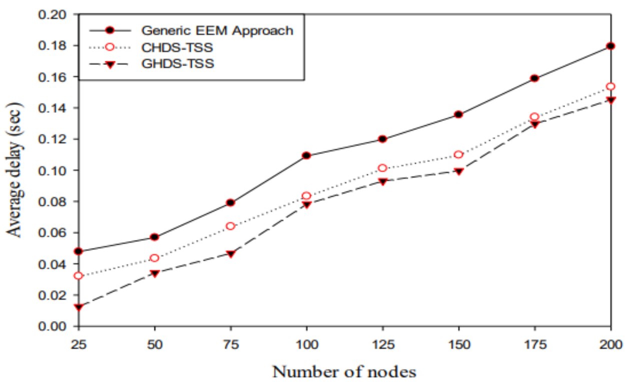

Fig. 4. Average delay for Generic EEM Approach, CHDS-TSS and GHDS-TSS under different number of nodes.

Power use.

When compared to the general EEM method, CHDS-TSS outperforms it in terms of average energy usage in simulations. The reason for this is that it relies on an improved EM clustering algorithm that incorporates optimum sink movement. This method aims to decrease the distance between sensor nodes and cluster heads, as well as between cluster heads and the sink, thus drastically reducing energy consumption at the sensor nodes. The reason for this behavior is because we have made every effort to maximize the benefits of communication distance.

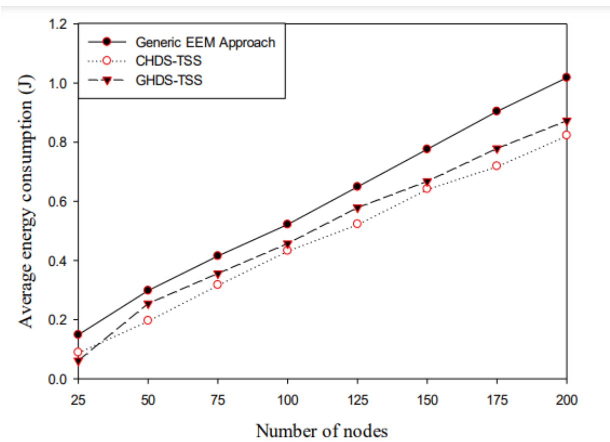

Fig. 5. Average energy consumption for Generic EEM Approach, CHDS-TSS and GHDS-TSS under different number of nodes.

Distribution ratio of packets

The suggested approaches outperformed the typical EEM method in terms of PDR. Figure 5 displays the findings for PDR. When compared to CHDS-TSS and the general EEM method, the results show that GHDS-TSS increases PDR. The decrease in data traffic, congestion, and computational complexity between the sink and head nodes of the network is the source of this increase in PDR. GHDS-TSS controls the buffer overflow issue and decreases the maximum likelihood of packet loss by reaching the target at the proper time.

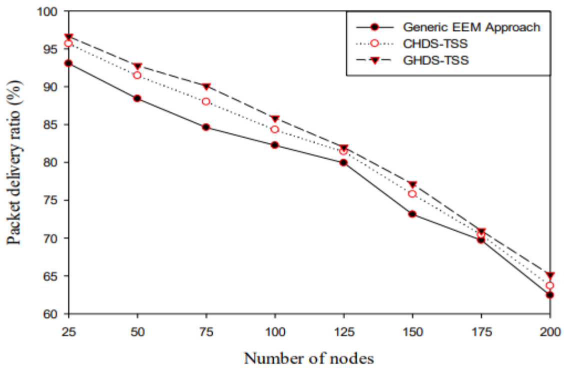

Fig. 6. Packet delivery ratio for Generic EEM Approach, CHDS-TSS and GHDS-TSS under different number of nodes.

Network durability

Several factors, including long-distance transmission, uneven load distribution, network obsoleting, and so on, impact the network's longevity. Figure 6 shows the network lifetime outcomes for several network scenarios; the suggested CHDS-TSS performs better than the others because of its balanced energy consumption across sensor nodes, its decrease in latency, and the number of hops it uses.

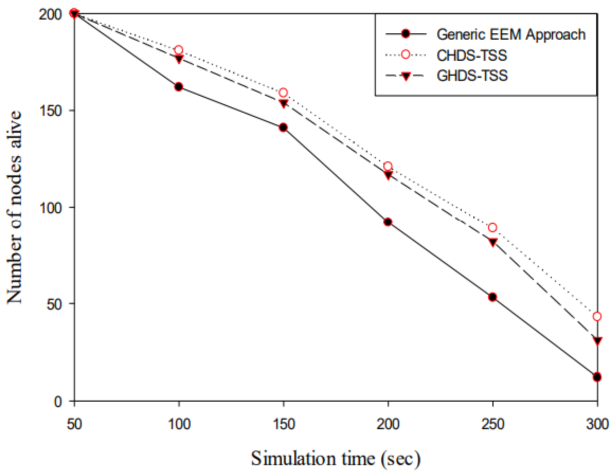

Fig. 7. Network lifetime for Generic EEM Approach, CHDS-TSS and GHDS-TSS under different simulation time.

CONCLUSION

The potential uses of WSNs are vast and diverse. The use of these networks to Big Data demonstrates their capacity to fulfill specific needs by overcoming inherent limitations. In Big Data, we anticipate a diverse dataset that is continually generated by the IoT, to the point that it surpasses the capabilities of conventional systems in terms of collection, management, and processing.

REFERENCES

1. Khan, A. W., Abdullah, A. H., Anisi, M. H., & Bangash, J. I. (2014). A Comprehensive Study of Data Collection Schemes Using Mobile Sinks in Wireless Sensor Networks. Sensors, 14(2), 2510-2548. https://doi.org/10.3390/s140202510

2. V. laxmi, "An Efficient Way to Gather Big data in WSN using Mobile Sink Routing," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, No. 8, pp. 43-45, August 2015.

3. A. Kaswan, K. Nitesh and P.K. Jana, "Energy Efficient Path Selection for Mobile Sink and Data Gathering in Wireless Sensor Networks," International Journal of Electronics and Communications, Vol. 73, No. 1, pp. 110-118, March 2017.

4. L. D. Xu, W. He, and S. Li, "Internet of Things in industries: A survey," IEEE Trans. Ind. Informat., vol. 10, no. 4, pp. 2233–2243, Nov. 2014, doi: 10.1109/TII.2014.2300753.

5. S. K. Singh, M. P. Singh, and D. K. Singh, "Routing protocols in wireless sensor networks—A survey," Int. J. Comput. Sci. Eng. Surv., vol. 1, no. 2, pp. 63–83, Nov. 2010, doi: 10.5121/ijcses.2010.1206.

6. A. Mahlous, "SCMC: An efficient scheme for minimizing energy in WSNs using a set cover approach," Future Internet, vol. 9, no. 4, p. 95, Dec. 2017, doi: 10.3390/fi9040095.