Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.755230

EGA-DAS: ENHANCED GENETIC ALGORITHM WITH EFFECTIVE DATA AGGREGATION SCHEME FOR AVOIDANCE OF DATA FLOODING IN WSN

L.Revathi

Research scholar, Department of Computer Applications, Dr.MGR Educational and Research Institute, Chennai, Tamil Nadu. reddy.revathi24@gmail.com

Dr.S.Kevin Andrews

Professor, Department of Computer Applications, Dr.MGR Educational and Research Institute, Chennai, Tamil Nadu. kevin.mca@drmgrdu.ac.in

Abstract

In Wireless Sensor Networks (WSNs), efficient routing is essential for maintaining network reliability and optimizing energy consumption. This paper introduces Enhanced Genetic Algorithm with Data Aggregation scheme (EGA-DAS) method a novel route discovery approach combining Electro-Magnetism with Enhanced Genetic Algorithm (EMEGA) to enhance routing efficiency. EMEGAutilizes electromagnetic principles for node attraction and disgust, simulating the optimization process similar to genetic algorithms. By iteratively refining routes based on energy consumption and network conditions, EMEGA aims to discover paths that minimize transmission costs and ensure robust communication. Performance evaluations demonstrate EMEGA's capability to achieve superior routing efficiency compared to traditional methods, offering promising advancements for reliable and energy-efficient WSN deployments. Extensive simulations and performance evaluations demonstrate that the proposed data aggregation scheme significantly outperforms existing methods in terms of energy efficiency, delivery delay, and data accuracy. This study contributes to the advancement of WSN technologies by offering a robust solution to the challenges of data flooding and delivery delay, making the way for more efficient and sustainable network operations.

Keywords:EGA-DAS, Data Aggregation,Adaptive Threshold-Based Filtering, Genetic Algorithm, Wireless Sensor Networks

I. Introduction

Many sensor nodes spread out over a network keep tabs on the surrounding environment and relay that information to a single "sink" node for analysis in a WSN [1]. Among the many uses for these networks are healthcare, industrial automation, environmental monitoring, and military surveillance [2]. Efficient data transmission, managing energy consumption, and guaranteeing reliable communication despite sensor node resource limits are the key problems in WSNs [3]. In addition to making sensor operating easier, modern WSNs operate in a bidirectional fashion [4]. When compared to a conventional cable network in uncertain circumstances, a WSN is more trustworthy since its sensor nodes can communicate and process data [5]. To minimize energy consumption and maximize the network lifetime in WSNs, effective data aggregation plays a crucial role in reducing redundant data transmission [6]. huge amounts of data are often generated by the many sensor nodes that make up WSNs, which are spread out across a huge region to keep an eye on the environment [7]. Data flooding, higher energy usage, and, eventually, network congestion and failure, can result from ineffective data aggregation [8]. In dynamic settings, where network circumstances and data relevance might change quickly, traditional data aggregation strategies, while beneficial, frequently fail [9]. By using adaptive threshold-based data filtering techniques, this work presents a new data aggregation approach that aims to tackle these issues [10]. This adjusts the aggregation process in real-time depending on the significance of the data and the state of the network, ensuring that only meaningful and non-redundant data is sent to the sink node [11].

Network circumstances and application needs can change fast in dynamic settings, which WSNs commonly operate in [12]. Therefore, adaptive data aggregation methods that can react to these changes in real-time and guarantee optimum performance in different environments are critically needed [13–14]. Timely data transmission is critical for decision-making and responsiveness in time-sensitive applications, making delivery delay another significant measure [15-16]. In this study, we provide a data aggregation strategy that can improve delivery latency in WSNs and prevent data flooding [17]. For the purpose of avoiding the transmission of unnecessary or irrelevant data, our method integrates a new data aggregation technique with adaptive threshold-based data filtering mechanisms [18–19]. Our technique successfully reduces the likelihood of data flooding and network congestion by adapting the data aggregation process in real-time according to network circumstances and data importance [20-22]. In order to make sure that data is sent promptly and reliably, we also use a priority-based data forwarding technique [23–24].

The main contribution of the paper

- Adaptive Significance-Based Data Aggregation
- Priority-Based Data Forwarding
- Route discovery using Electro-Magnetism with Enhanced Genetic Algorithm

What follows is the outline for the rest of the article. A wide range of data aggregation methodologies are covered by several authors in Section 2. In Section 3, we can see the EGA-DAS model. The investigation's findings are summarized in Section 4. Finally, Section 5 explore into a review of the outcome and potential future endeavors.

1.1 Motivation of the paper

Improving routing efficiency and minimizing energy usage are two major issues in Wireless Sensor Networks, which is why this study is dedicated to solving these problems. For WSNs to have dependable communication and a long lifetime, efficient routing is crucial. This work intends to considerably increase routing efficiency via the creative application of electromagnetic principles and genetic algorithms by developing the Electro-Magnetism with Enhanced Genetic Algorithm route finding technique. The main objective is to optimize network resilience while reducing transmission costs via the implementation of dynamic route adaptation in response to real-time energy and network circumstances. By providing a solid solution that guarantees more dependable, energy-efficient, and environmentally friendly network operations, this study aids in the advancement of WSN technology.

II. Background study

Chen, Y. et al. [1] the author presented a method for collecting data from smart meters in this research. Thanks to the suggested method, smart meters can now provide data in several dimensions. This allows the utility supplier to analyze the data more thoroughly, learn about data variation, do one-way analysis of variance, and more. This study's signature system outperforms comparable research in terms of efficiency, leading to quicker message verification on both the aggregator and utility provider sides.

Cui, J. et al. [3] the author fix several mistakes and fix the security holes in existing data aggregation methods, and the author successfully propose a method to secure the CDAMA scheme's data. The author also provides a low-power, secure data aggregation method that can handle large-scale WSNs. For end-to-end data secrecy, the author use the OU homomorphic encryption method. For in-network false data filtering, the author use MAC. For end-to-end data integrity, the author use the homomorphic MAC algorithm.

Gai, N. et al. [5] the author presented a privacy-preserving smart grid data aggregation approach in this research. The author lessened the computing load on smart meters that take part in data aggregation duties, taking into account their restricted computational capability. The approach accomplishes the privacy-preserving smart grid data aggregation that satisfies the LDP by developing a unique algorithm for data discretization and a random response mechanism. In the smart grid data aggregation scenario, the author take into account additional specific cases that allow these authors scheme to handle the unique circumstance of broad data range.

Gupta, S., &Snigdh, I.[8] these authors research suggests a technique for aggregating data packets in LoRa networks. The author has evaluated the energy consumption of several methods for data packet transmission and aggregation, including Me-Cat, traditional LoRa, and clustering LoRa. These authors suggested approach beats the other two algorithms by a significant margin. By comparing the amount of data packets delivered by each method, the author discovered that traditional LoRa and data aggregation techniques send out packets with comparable sizes.

Lu, R. et al. [11] For the Internet of Things (IoT) that makes use of fog computing, the author provide LPDA, a lightweight data aggregation approach that protects users' privacy. By placing the fog device at the periphery of the network, LPDA was able to do more than just prevent malicious actors from injecting misleading data; it can also facilitate fault tolerance and efficiently combine data from several hybrid IoT devices. These authors defined security model verifies that the suggested LPDA technique was safe via extensive security testing, including increased differential privacy testing.

Shen, H. et al. [15] an effective and privacy-preserving technique for the aggregation of power consumption cube-data in smart grids was presented in this research. The method can collect multi-dimensional data from different residential regions for many consumers at different levels of detail. Flexible electricity regulation, totaling all residential areas' power consumption across all dimensions, and regulation of electricity for each residential area across all dimensions were all within the control center's purview.

Wang, X. et al. [18] these authors provided VSDA, a data aggregation protocol, to solve the problems that arise when a WSN was scaled up. In the same way as CS-based systems encapsulate sensor node raw data in a weight vector, VSDA uses the same vectors to construct a measurement matrix at the sink node, where sensor data was decoded.

Zhang, J., & Dong, C.[21] In order to aggregate data in a way that protects users' privacy, there were two major challenges: security and privacy. Due to the aggregator's status as an honest-but-curious entity, the majority of current data aggregation systems have inadequate security. The authors provide a new, lightweight privacy-preserving data aggregation technique that protects users' data against modification and deletion attempts by malevolent aggregators, and they do so by using their symmetric homomorphic encryption approach.

Table 1: Survey on Data Aggregation Schemes

Author	Year	Methodology	Advantage	Limitation
Chen, Y. et al.	2019	Homomorphic-based aggregation	Enhances data security and privacy	High computational overhead
Cui, J. et al.	2018	End-to-end confidentiality and integrity	Ensures data integrity throughout transmission	Increased communication overhead
Gai, N. et al.	2022	Differential privacy	Protects individual data while enabling useful analysis	Complexity in parameter tuning
Shen, X. et al.	2020	Privacy-preserving aggregation	Supports dynamic group scenarios	Potential reduction in aggregation efficiency
Zhang, X.	2022	Blockchain-based aggregation	Provides robustness against tampering and deletion attacks	Blockchain overhead and scalability issues

2.1 Problem definition

To optimize energy usage and reduce duplicate data transmission, data aggregation plays a crucial role in WSNs. Nevertheless, current methods often encounter difficulties including susceptibility to security risks, wasteful energy use, and insufficient scalability. For example, existing methods can not be strong enough to withstand data manipulation or can't handle with the ever-growing number of network installations. There is already a lot of complexity in the landscape due to privacy issues and the need for effective data aggregation in smart grid applications. In Genetic algorithms has do not scale well with complexity so, in this paper has enhanced Genetic algorithm.In response to these issues, the authors of this study provide EGA-DAS, a new data aggregation technique that can improve network efficiency, delay delivery, and avoid data flooding.

III. Materials and methods

In this section, we detail the proposed method for enhancing routing efficiency in Wireless Sensor Networks through the Electro-Magnetism with Enhanced Genetic Algorithm (EMEGA). EMEGA combines electromagnetic principles with genetic algorithms to optimize route discovery by utilizing node attraction and repulsion dynamics.

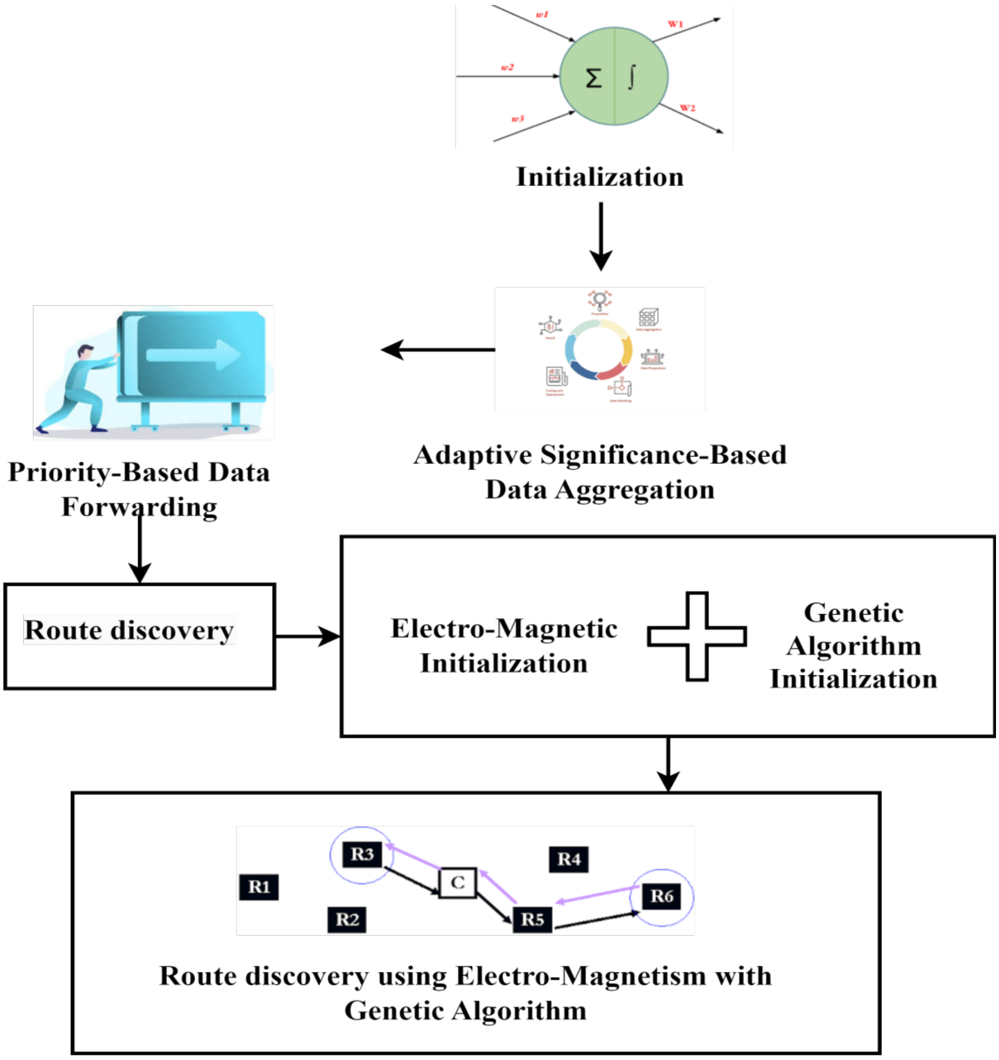

Figure 1: EGA-DAS workflow architecture

3.1 Network model

In order to test how well our suggested data aggregation method works in WSNs, we set up a network model with N sensor nodes spread out throughout a monitoring region. Nodes 1, 2,..., N in the sensor network are all tasked with collecting data about their surroundings and relaying it to node S, the sink node. The following parts and equations make up the network model.

a. Sensor Node Model

Each sensor node i generates data packets at regular intervals. The data packet size is denoted by d_i bits. The energy consumption for transmitting a data packet from node i to node j over a distance d_{ij} is given by:

$E_{tx}(i,j) = d_i.\left(E_{elec} + \epsilon_{amp}.d_{ij}^2\right)$------------ (1)

where:

E_{elec} is bit-level power consumption of the transmitter and receiver circuits

ϵ_{amp} is the energy consumption per bit per square meter for the transmission amplifier.

d_{ij} is the distance between nodes i and j.

The energy consumption for receiving a data packet is given by:

$E_{rx}(i) = d_i . E_{elec}$-------------- (2)

The proposed data aggregation scheme involves two main components: a novel data aggregation algorithm and adaptive threshold-based data filtering.

The aggregation function f_{agg} combines data packets from multiple sensor nodes into a single aggregated packet to reduce redundancy:

$d_{agg} = f_{agg}(d_1, d_2, \dots, d_n)$------------- (3)

Each sensor node applies a threshold-based filtering mechanism to determine the significance of its data before transmission. The significance S_i of data from node i is evaluated using:

$S_i = \frac{d_i - \mu}{\sigma}$-------------- (4)

where:

- μ is the mean of the data values
- σ is the standard deviation of the data values

3.2 Adaptive Significance-Based Data Aggregation

Wireless sensor networks can benefit from the Adaptive Significance-Based Data Aggregation (ASDA) method, which is specifically developed to enhance data transmission. Reducing redundancy and energy usage, it combines and filters key data. In response to changes in the network and the importance of the data, the algorithm adapts the filtering threshold on the fly. To create a single representative packet from data packets received by several sensor nodes, ASDA uses a number of aggregation methods, including weighted averaging and averaging. This method improves the performance and longevity of the network by guaranteeing the timely, accurate, and efficient transfer of data.

As mentioned earlier, it is computationally difficult to compute the aggregation levels, cluster topologies, and feature coefficients simultaneously, which might result in considerable prediction errors. Our novel approach can do the following: (a) find the best aggregation level for each feature; (b) find the underlying cluster structure of the SKUs with respect to each feature; and (c) consistently estimate the feature coefficients. Our data-driven methodology not only generates a trustworthy demand prediction, but it also has the potential to efficiently and effectively fulfill all three of these goals.

We start by examining a (basic) subset of the GLM, where all characteristics are at the SKU level, as a particular instance. Here, we can express the data-generating process as

$Y_{i,j} = G\left(\sum_{l \in D} X_{i,j}^{l} b_{i,l}\right) + \epsilon_{i,j}$------ (5)

$i = 1, \dots, n and j = 1, \dots, m.$--------- (6)

We get $b_{i,l}$ for l ∈ Ds, $b_{i,l}$ βn i,l for $\epsilon_{i,j}$, and $b_{i,l}$,l for l ∈ Dc by comparing the model specifications in (5) and (6). Since each item is fitted in a decentralized method, we call Model (6) the decentralized model. Iterative reweighted least squares is often used to estimate the decentralized model. For typical GLMs like logistic and linear regression, we assume that each item has a well-defined decentralized model with a distinct MLE solution. It is possible to estimate the decentralized model by breaking it down into individual item-specific estimates.

$b_i \in arg \sum_{j=1}^{m} log L_i \left(y_{i,j}, x_{i,j}\right)$------ (7)

In this case, $(y_{i,j}, x_{i,j})$ represents the likelihood function associated with the data $(y_{i,j}, x_{i,j})$ and the coefficient vector bi ∈ Rd, and H(•): In the generalized linear model (GLM), the normalization mapping from R to H is indefinitely differentiable.

For details, see Online Appendix A. $y_{i,j}, x_{i,j}$. Our decentralized estimator is the set $(y_{i,j}, x_{i,j})$. To make this reliance obvious, we parameterize the estimators with the sample size m throughout this study. As an example, a decentralized estimator with a sample size of m is represented as $(y_{i,j}, x_{i,j})$. The Fisher information matrix is defined in terms of the decentralized item i model as well.

$I_i(b_i) := -E[\nabla_2 log L_i(y_{i,j}, X_{i,j})]$------ (8)

within the context of $(y_{i,j}, X_{i,j})$, with ∇_2 representing the Hessian operator and the expectation being measured. To provide the groundwork for our further analysis, we first demonstrate the decentralized estimator b's following consistency and normalcy properties.

3.3 Priority-Based Data Forwarding

Wireless sensor networks and other communication networks use priority-based data forwarding to guarantee efficient and timely data transfer. It uses factors like data packet importance and time sensitivity to determine their priority levels. In order to fulfill application needs like critical event reporting or real-time data transmission, higher-priority packets are transmitted with precedence over lower-priority ones. By prioritizing the transmission of critical data and efficiently managing network resources, this technique improves bandwidth consumption and boosts overall system performance.

The order of importance is based on the rate of energy harvesting and surplus power. The priority P is determined by the nodes by considering the remaining power and the energy-harvesting rate. A lower energy-harvesting rate allows for more frugal utilization of leftover power, regardless of how much of it there is. In this scenario, a node with a large residual power but a slow energy harvesting rate can be disabled since its battery power is soon depleted. A high energy collecting rate, on the other hand, will charge the node enough during each data transmission period to make it a preferred relay node even if the leftover power is little. As seen in (1), the priority of the node is determined by T0, which takes into account both the residual power and the energy-harvesting rate.

$P_i = E_{r,i} + R_{h,i}T_0$ --------- (9)

in which Pi takes precedence, Rh,i is the rate of energy harvesting for the i-th node, while Er,i is the residual power. Te priorità The estimated remaining power of the node at the next transmission time after the sleep period is represented by pi. Until their power is charged over the minimum power, nodes will remain in the sleep state since they do not have enough energy to transmit data, even if they win access to the channel via contention.

The priority of data packets in each CR's transmission queue determines which CR's forwarder is used in tuning-based forwarder selection. We have developed a cost function that considers both stability and latency. From the perspective of CF_i, let $CF_{i,j}$ represent the cost of using downstream $S_{i,j}$ as the forwarder. Two weights, α and H, are given to the chosen metrics in order of relative importance. This is how the cost function is calculated:

$CF_{i,j} = a \times D_j + \beta \times S_{i,j}$----------- (10)

$a + \beta = 1$------------ (11)

If CR ci want to transmit an HP packet, it is ideal for data transfer to take place with minimal delay. Therefore, it determines the best forwarder CR based on its lowest cost by using Eq.10, which is supplemented by adding the weight related to delay, to calculate the cost value $CF_{i,j}$ for each ci $S_{i,j}$.

3.4 Route discovery using Electro-Magnetism with Genetic Algorithm

To aid in the finding of routes in communication networks, particularly WSN, the EMEGA combines electromagnetism with genetic algorithms. Like electromagnetism, EMEGA depicts nodes as charged particles with attracting and repulsive forces depending on their placements. Afterwards, genetic algorithms optimize possible routes between nodes repeatedly based on fitness factors such path length, energy efficiency, or network congestion. This mixed method improves communication performance and reliability by balancing exploration and exploitation, which allows for efficient and adaptable route finding under changing network circumstances.

In order to identify circles in data, this research employs an EM-like method. To optimize global multi-modal functions, the EM algorithm employs a straightforward and simple population-based search strategy. Its use of mutation and crossover operators to discover potentially useful sections is inferior to that of a genetic algorithm. Physical principles provide the basis of the algorithm.

A subset of optimization problems involving limited variables can be addressed using the EM method by formulating

$min f(x)$ --------------- (12)

$x \in [l, u]$ --------------- (13)

Initialization, local search, computation, and movement are the four primary steps of an algorithm similar to EM. These are characterized as

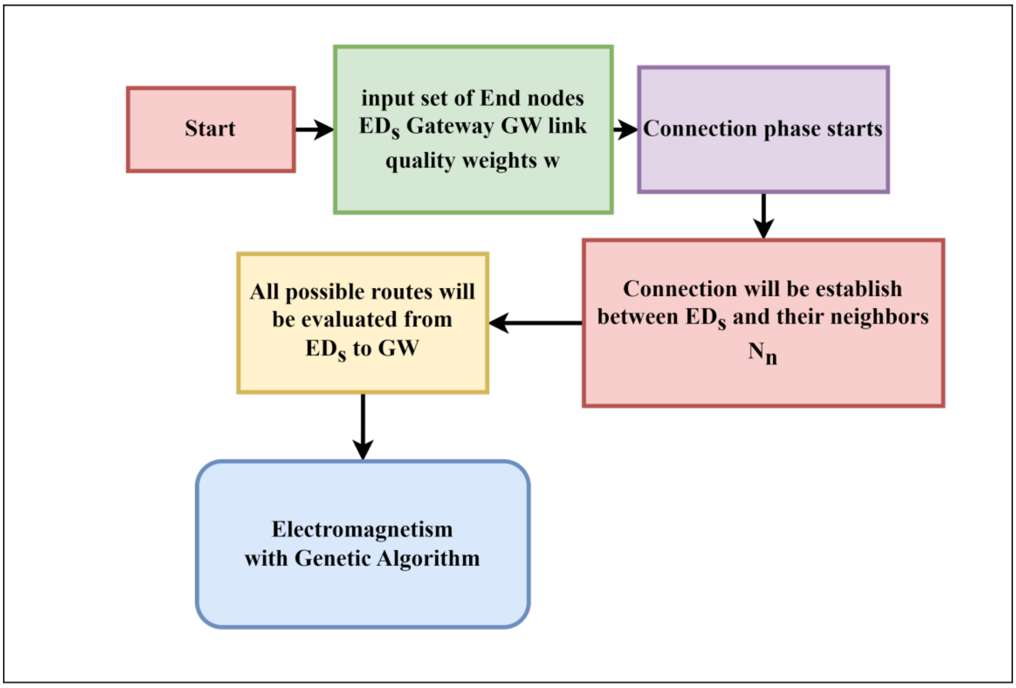

Figure 2: EMEGA flowchart

Each gene's position on a chromosome stands in for a node in the network; for example, gene 3 is located on node 3. The value of the gene at each point on a node indicates the next hop node. The starting population is a group of chromosomes. The current approach uses a randomly generated set of functional chromosomes as the genetic algorithms starting population. All of the chromosome's pathways must be free of rings for the chromosome to be considered genuine, and each gene's value must match to a legitimate next hop on the related relay.

$Max(z) = \sum_{i=1}^{m} Cost(i,j)$--------------- (14)

Where$i$, j denote the i-th and j-th relay node and $Cost(i,j)$ denotes the cost of data transmission from node i to no

Algorithm 1: Electro-Magnetism with Enhanced Genetic Algorithm
Input: Number of nodes, their positions, connectivity, and transmission costs **Begin:** Similar to EM (Expectation-Maximization) but adapted for detecting circles in data. Focuses on population-based search without crossover and mutation, utilizing physical principles for optimization. Handles bounded variables (constraints) in optimization tasks. Genetic Algorithm (GA) for Route Optimization Represent sequences of nodes (paths) in the network. Random generation of initial population. Ensures paths are ring-free and valid Maximizes the sum of transmission costs between relay nodes$\left(Cost(i,j)\right)$. $(z) = \sum_{i=1}^{m} Cost(i,j)$ i,jDenote relay nodes. $Cost(i,j)$ represents the cost of data transmission from node i to node j **End** Sequence of nodes representing the optimal path(s) between specified nodes in the network.

The EMEGA combines physical principles of electro-magnetism for node optimization and genetic algorithm techniques for route discovery in WSNs. It integrates EM-like interactions to optimize node positions and genetic algorithm approaches to refine optimal routes based on transmission costs. This hybrid approach aims to enhance routing efficiency by dynamically adjusting node positions and selecting the most cost-effective paths, ensuring robust communication and energy efficiency in WSN deployments.

IV. Results and discussion

This section presents the simulation results of the EGA-DAS protocol and evaluates it against the existing 6LowPAN-Aggr, LORA, EEDAMand F-LEACHscheduling protocols using Network Simulator (NS-3). The simulation is conducted in two different scenarios, focusing on sensing reliability with variations and network density. The simulation encompasses a set of parameters, which can be found in Table 2.

Table 2: Simulation Parameters

Parameters	Values
No. of sensor nodes	250

Simulation area	1000×1000m^2
Sensing length	50m
Routing protocol	6LoWPAN
Queue type	CMUPriQueue
Packet size	300bits
Buffer length	65 packets
Initial node energy	70J
MAC type	MAC/802.11
Simulation time	65ms

In the context of proposed method can evaluate several important performance metrics including Packet Delivery Ratio (PDR), Packet Loss, Network Lifetime, Energy Consumption, Average Delay, Average Throughput, and Communication Overhead.

Table 3: Comparison of Average Throughput in mbps

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	0.60	0.66	0.73	0.80	0.93
100	0.55	0.62	0.65	0.71	0.86
150	0.47	0.51	0.56	0.65	0.75
200	0.40	0.43	0.50	0.52	0.64
250	0.35	0.38	0.42	0.50	0.60

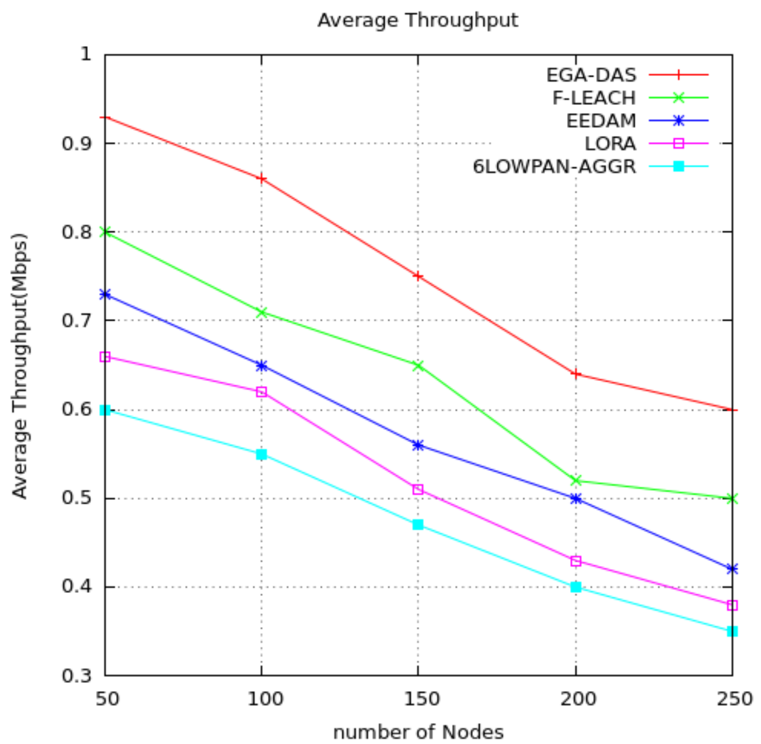

Figure 3: Comparison chart of Average Throughput in mbps

The table 3 and figure 3 shows provided various routing protocols in terms of energy efficiency across different node counts shows notable variations. At 50 nodes, 6LowPAN-Aggr achieves an energy efficiency of 0.60, LORA 0.66, EEDAM 0.73, F-LEACH 0.80, and EGA-DAS 0.93. As the node count increases to 100, 6LowPAN-Aggr's efficiency drops to 0.55, LORA to 0.62, EEDAM to 0.65, F-LEACH to 0.71, and EGA-DAS to 0.86. At 150 nodes, the trend continues with 6LowPAN-Aggr at 0.47, LORA at 0.51, EEDAM at 0.56, F-LEACH at 0.65, and EGA-DAS at 0.75. For 200 nodes, 6LowPAN-Aggr further declines to 0.40, LORA to 0.43, EEDAM to 0.50, F-LEACH to 0.52, and EGA-DAS to 0.64. Finally, at 250 nodes, 6LowPAN-Aggr has an efficiency of 0.35, LORA 0.38, EEDAM 0.42, F-LEACH 0.50, and EGA-DAS 0.60. Overall, EGA-DAS consistently exhibits the highest energy efficiency across all node counts, followed by F-LEACH and EEDAM, whereas 6LowPAN-Aggr and LORA show lower performance, particularly as the node count increases.

Table 4: Comparison of PDR in %

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	86	87	90	92	97
100	81	82	85	88	91
150	75	76	80	83	88
200	71	73	78	80	86
250	69	71	74	77	82

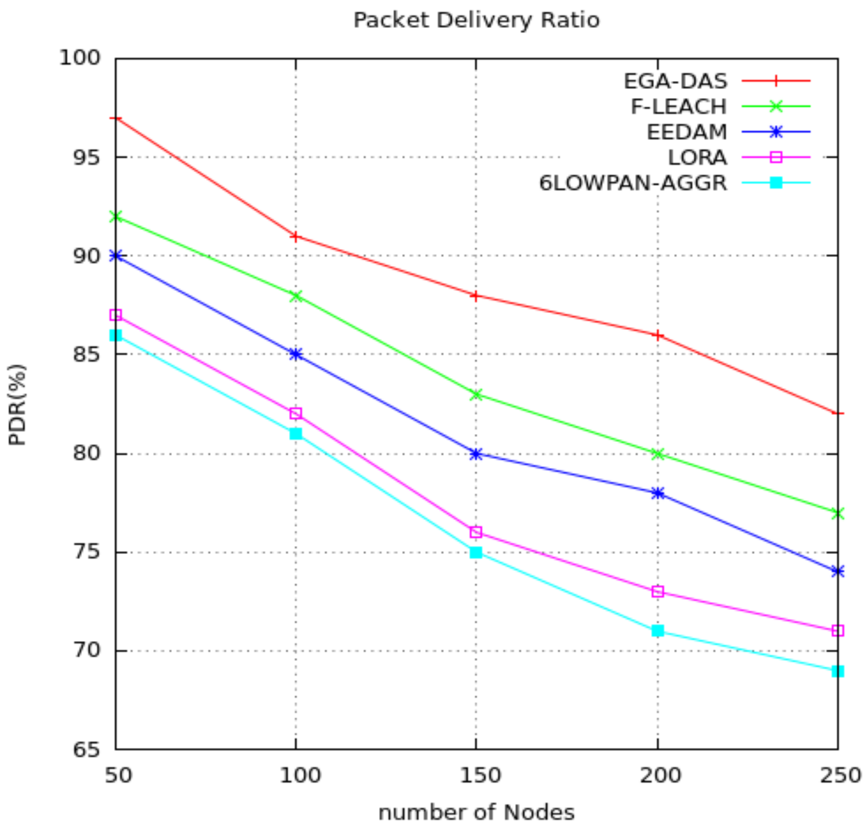

Figure 4:Comparison chart of PDR in %

The table 4 and figure 4 presents the comparison of various routing protocols in terms of their performance across different node counts reveals distinct patterns. At 50 nodes, 6LowPAN-Aggr has a performance value of 86, LORA 87, EEDAM 90, F-LEACH 92, and EGA-DAS 97. As the node count increases to 100, 6LowPAN-Aggr's performance drops to 81, LORA to 82, EEDAM to 85, F-LEACH to 88, and EGA-DAS to 91. With 150 nodes, 6LowPAN-Aggr achieves a performance value of 75, LORA 76, EEDAM 80, F-LEACH 83, and EGA-DAS 88. At 200 nodes, the values are 71 for 6LowPAN-Aggr, 73 for LORA, 78 for EEDAM, 80 for F-LEACH, and 86 for EGA-DAS. Finally, at 250 nodes, 6LowPAN-Aggr has a performance value of 69, LORA 71, EEDAM 74, F-LEACH 77, and EGA-DAS 82. Overall, EGA-DAS consistently demonstrates the highest performance across all node counts, followed by F-LEACH and EEDAM, while 6LowPAN-Aggr and LORA show comparatively lower performance, particularly as the node count increases.

Table 5: Comparison of Packet Loss in percentage

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	14	12	10	7	3
100	19	18	15	11	9
150	25	23	20	17	12
200	29	27	22	20	14
250	31	29	26	23	18

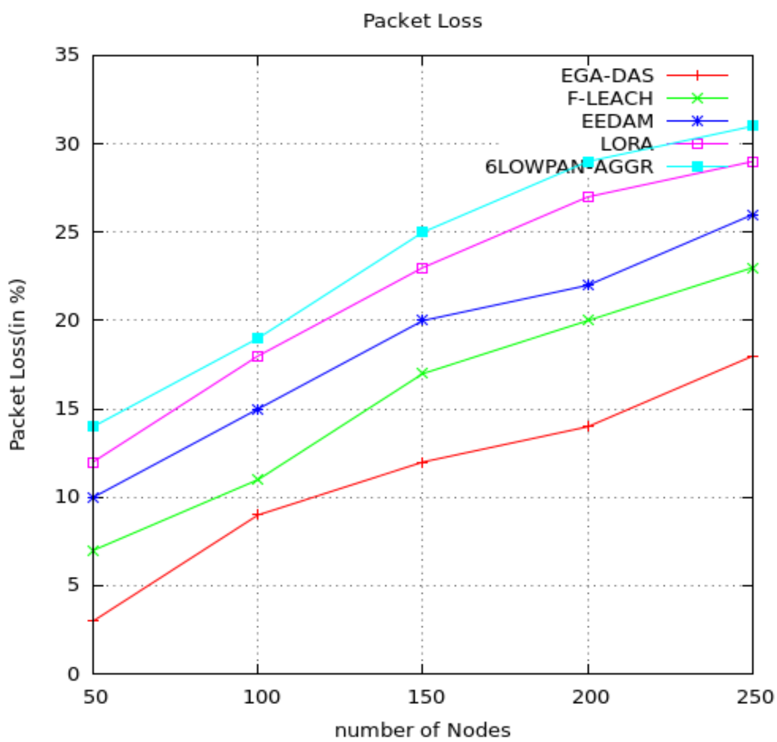

Figure 5: Comparison chart of Packet Loss in %

The table 5 and figure 5 presents the performance of different routing protocols in terms of the delay across varying node counts is analyzed, showing a trend of increasing delay with higher node counts. At 50 nodes, 6LowPAN-Aggr exhibits a delay of 14 ms, LORA 12 ms, EEDAM 10 ms, F-LEACH 7 ms, and EGA-DAS 3 ms. When the node count increases to 100, the delay for 6LowPAN-Aggr rises to 19 ms, LORA to 18 ms, EEDAM to 15 ms, F-LEACH to 11 ms, and EGA-DAS to 9 ms. At 150 nodes, 6LowPAN-Aggr experiences a delay of 25 ms, LORA 23 ms, EEDAM 20 ms, F-LEACH 17 ms, and EGA-DAS 12 ms. For 200 nodes, the delays are 29 ms for 6LowPAN-Aggr, 27 ms for LORA, 22 ms for EEDAM, 20 ms for F-LEACH, and 14 ms for EGA-DAS. Finally, at 250 nodes, 6LowPAN-Aggr has a delay of 31 ms, LORA 29 ms, EEDAM 26 ms, F-LEACH 23 ms, and EGA-DAS 18 ms. Overall, EGA-DAS consistently demonstrates the lowest delay across all node counts, followed by F-LEACH and EEDAM, while 6LowPAN-Aggr and LORA show relatively higher delays, especially as the node count increases.

Table 6: Comparison of Network Lifetime in m.seconds

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	100	100	100	100	100
100	90	92	95	98	99
150	84	87	90	93	97
200	80	85	88	90	96
250	77	82	84	87	92

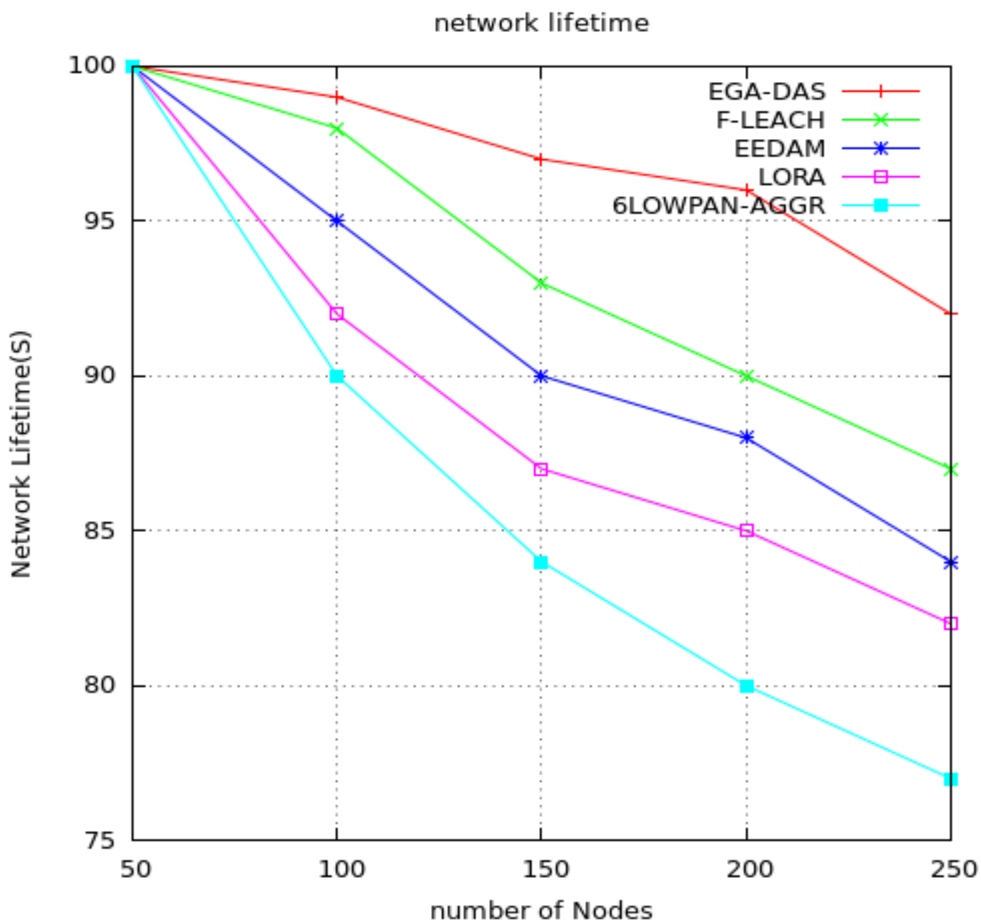

Figure 6: Comparison chart of Network Lifetime in s

The table 6 and figure 6 shows the comparison of various routing protocols in terms of their reliability across different node counts indicates that all protocols perform consistently well at lower node counts, with some differentiation as the node count increases. At 50 nodes, all protocols (6LowPAN-Aggr, LORA, EEDAM, F-LEACH, and EGA-DAS) achieve a reliability score of 100. When the node count increases to 100, the reliability for 6LowPAN-Aggr drops to 90, LORA to 92, EEDAM to 95, F-LEACH to 98, and EGA-DAS to 99. At 150 nodes, 6LowPAN-Aggr scores 84, LORA 87, EEDAM 90, F-LEACH 93, and EGA-DAS 97. For 200 nodes, the reliability values are 80 for 6LowPAN-Aggr, 85 for LORA, 88 for EEDAM, 90 for F-LEACH, and 96 for EGA-DAS. Finally, at 250 nodes, 6LowPAN-Aggr has a reliability score of 77, LORA 82, EEDAM 84, F-LEACH 87, and EGA-DAS 92. Overall, EGA-DAS consistently exhibits the highest reliability across all node counts, followed closely by F-LEACH and EEDAM, whereas 6LowPAN-Aggr and LORA show comparatively lower reliability, particularly as the node count increases.

Table 7: Comparison of Energy Consumption in Joules

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	0.76	0.64	0.55	0.47	0.23
100	0.89	0.73	0.64	0.50	0.35
150	0.97	0.89	0.85	0.72	0.47
200	1.44	1.23	0.99	0.85	0.58
250	1.98	1.45	1.13	0.98	0.63

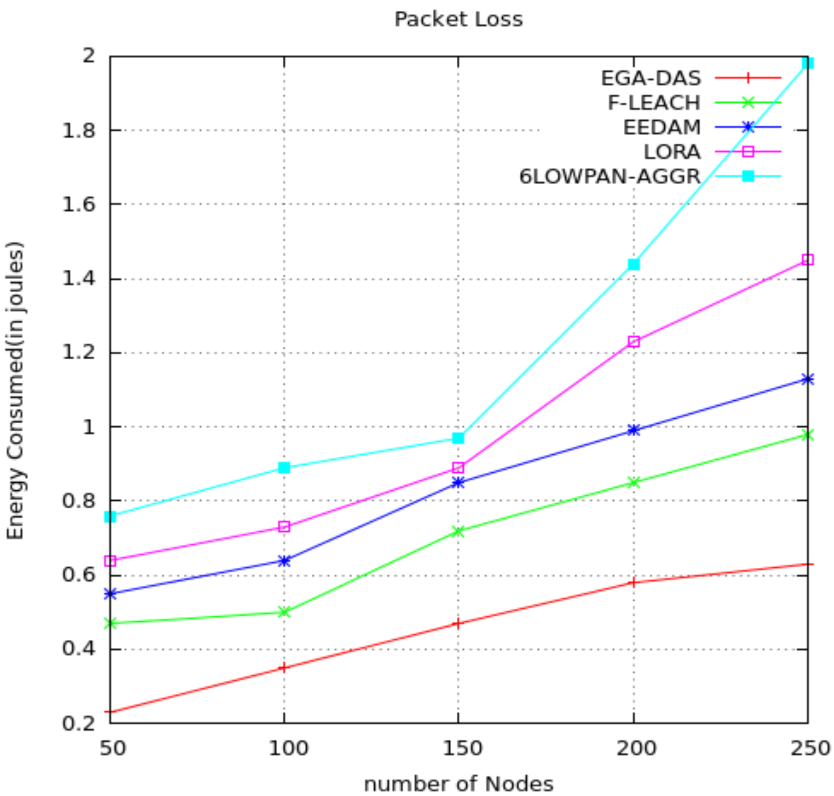

Figure 7: Comparison chart of Energy Consumption in J

The table 7 and figure 7 presents the comparison of energy consumption across various routing protocols for different node counts shows a clear trend of increasing energy usage with higher node counts, with notable differences between the protocols. At 50 nodes, 6LowPAN-Aggr has an energy consumption of 0.76 J, LORA 0.64 J, EEDAM 0.55 J, F-LEACH 0.47 J, and EGA-DAS 0.23 J. When the node count increases to 100, the energy consumption rises to 0.89 J for 6LowPAN-Aggr, 0.73 J for LORA, 0.64 J for EEDAM, 0.50 J for F-LEACH, and 0.35 J for EGA-DAS. At 150 nodes, 6LowPAN-Aggr consumes 0.97 J, LORA 0.89 J, EEDAM 0.85 J, F-LEACH 0.72 J, and EGA-DAS 0.47 J. For 200 nodes, the energy consumption values are 1.44 J for 6LowPAN-Aggr, 1.23 J for LORA, 0.99 J for EEDAM, 0.85 J for F-LEACH, and 0.58 J for EGA-DAS. Finally, at 250 nodes, 6LowPAN-Aggr has the highest energy consumption of 1.98 J, followed by LORA with 1.45 J, EEDAM with 1.13 J, F-LEACH with 0.98 J, and EGA-DAS with the lowest at 0.63 J. Overall, EGA-DAS consistently shows the lowest energy consumption across all node counts, indicating its superior efficiency, while 6LowPAN-Aggr and LORA exhibit the highest energy consumption, particularly as the node count increases.

Table 8: Comparison of Average Delay in ms

Node Count	6LowPAN-Aggr	LORA	EEDAM	F-LEACH	EGA-DAS
50	8.35	7.58	6.34	4.23	2.012
100	10.33	8.21	7.29	5.97	3.056
150	12.96	9.33	8.11	6.67	4.37
200	14.77	11.54	9.12	7.55	5.66
250	15.92	12.58	10.74	8.06	6.78

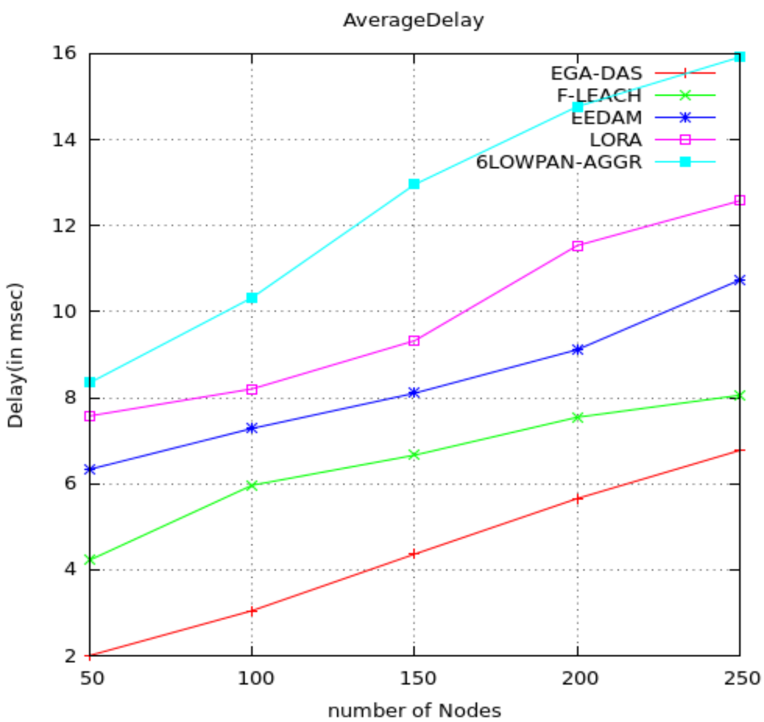

Figure 8: Comparison chart of Average Delay in ms

The table 8 and figure 8 displays the comparison of average delay across various routing protocols for different node counts highlights distinct variations in their performance. At 50 nodes, 6LowPAN-Aggr experiences an average delay of 8.35 ms, LORA 7.58 ms, EEDAM 6.34 ms, F-LEACH 4.23 ms, and EGA-DAS 2.012 ms. When the node count increases to 100, the delay rises to 10.33 ms for 6LowPAN-Aggr, 8.21 ms for LORA, 7.29 ms for EEDAM, 5.97 ms for F-LEACH, and 3.056 ms for EGA-DAS. At 150 nodes, 6LowPAN-Aggr has a delay of 12.96 ms, LORA 9.33 ms, EEDAM 8.11 ms, F-LEACH 6.67 ms, and EGA-DAS 4.37 ms. For 200 nodes, the delays are 14.77 ms for 6LowPAN-Aggr, 11.54 ms for LORA, 9.12 ms for EEDAM, 7.55 ms for F-LEACH, and 5.66 ms for EGA-DAS. Finally, at 250 nodes, 6LowPAN-Aggr exhibits the highest delay of 15.92 ms, followed by LORA with 12.58 ms, EEDAM with 10.74 ms, F-LEACH with 8.06 ms, and EGA-DAS with the lowest at 6.78 ms. Overall, EGA-DAS consistently demonstrates the lowest average delay across all node counts, indicating its efficiency in minimizing delay, whereas 6LowPAN-Aggr and LORA show the highest delays, particularly as the node count increases.

V. CONCLUSION

In Conclusion, EGA-DAS presents an innovative data aggregation scheme designed to address the critical issues of data flooding and delivery delay in WSNs. The proposed scheme integrates a novel data aggregation algorithm with adaptive threshold-based data filtering mechanisms to ensure the transmission of only significant and non-redundant data to the sink node. By dynamically adjusting the data aggregation process based on network conditions and data significance, the approach effectively mitigates the risks associated with data flooding and network congestion, leading to reduced energy consumption and enhanced network lifespan. Furthermore, the scheme employs a priority-based data forwarding strategy, which significantly improves delivery delay, ensuring timely and reliable data transmission for time-sensitive applications. The results highlight the scheme's potential to provide a robust solution to the prevalent challenges in WSNs, thereby contributing to the advancement and sustainability of WSN technologies. Future work will focus on further optimizing the adaptive mechanisms to handle a wider range of dynamic network conditions and application scenarios. Additionally, exploring the integration of machine learning techniques for predictive data aggregation and anomaly detection in WSNs presents an exciting avenue for enhancing the robustness and intelligence of the proposed scheme.

VI. References

1. Chen, Y., Martínez-Ortega, J. F., Castillejo, P., &López, L. (2019). A homomorphic-based multiple data aggregation scheme for smart grid. IEEE Sensors Journal, 19(10), 3921-3929.
2. Chen, Y., Martínez-Ortega, J. F., Castillejo, P., &López, L. (2019). An elliptic curve-based scalable data aggregation scheme for smart grid. IEEE Systems Journal, 14(2), 2066-2077.
3. Cui, J., Shao, L., Zhong, H., Xu, Y., & Liu, L. (2018). Data aggregation with end-to-end confidentiality and integrity for large-scale wireless sensor networks. Peer-to-Peer Networking and Applications, 11, 1022-1037.
4. Devi, V. S., Ravi, T., &Priya, S. B. (2020). Cluster based data aggregation scheme for latency and packet loss reduction in WSN. Computer Communications, 149, 36-43.

5. Gai, N., Xue, K., Zhu, B., Yang, J., Liu, J., & He, D. (2022). An efficient data aggregation scheme with local differential privacy in smart grid. Digital Communications and Networks, 8(3), 333-342.
6. Goyal, N., Dave, M., &Verma, A. K. (2020). SAPDA: secure authentication with protected data aggregation scheme for improving QoS in scalable and survivable UWSNs. Wireless Personal Communications, 113(1), 1-15.
7. Guan, Z., Zhang, Y., Zhu, L., Wu, L., & Yu, S. (2019). EFFECT: An efficient flexible privacy-preserving data aggregation scheme with authentication in smart grid. Science China Information Sciences, 62, 1-14.
8. Gupta, S., &Snigdh, I. (2022). Leveraging data aggregation algorithm in LoRa networks. The Journal of Supercomputing, 78(15), 16861-16875.
9. He, D., Zeadally, S., Wang, H., & Liu, Q. (2017). Lightweight data aggregation scheme against internal attackers in smart grid using elliptic curve cryptography. Wireless Communications and Mobile Computing, 2017(1), 3194845.
10. Li, A., Liu, W., Zeng, L., Fa, C., & Tan, Y. (2021). An efficient data aggregation scheme based on differentiated threshold configuring joint optimal relay selection in WSNs. IEEE access, 9, 19254-19269.
11. Lu, R., Heung, K., Lashkari, A. H., &Ghorbani, A. A. (2017). A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT. IEEE access, 5, 3302-3312.
12. Ming, Y., Li, Y., Zhao, Y., & Yang, P. (2022). Efficient Privacy-Preserving Data Aggregation Scheme with Fault Tolerance in Smart Grid. Security and Communication Networks, 2022(1), 5895176.
13. Ming, Y., Zhang, X., &Shen, X. (2019). Efficient privacy-preserving multi-dimensional data aggregation scheme in smart grid. IEEE Access, 7, 32907-32921.
14. Saleem, A., Khan, A., Malik, S. U. R., Pervaiz, H., Malik, H., Alam, M., & Jindal, A. (2019). FESDA: Fog-enabled secure data aggregation in smart grid IoT network. IEEE Internet of Things Journal, 7(7), 6132-6142.
15. Shen, H., Zhang, M., &Shen, J. (2017). Efficient privacy-preserving cube-data aggregation scheme for smart grids. IEEE Transactions on Information Forensics and Security, 12(6), 1369-1381.
16. Shen, X., Zhu, L., Xu, C., Sharif, K., & Lu, R. (2020). A privacy-preserving data aggregation scheme for dynamic groups in fog computing. Information Sciences, 514, 118-130.
17. Sajedi, S. N., Maadani, M., &NesariMoghadam, M. (2022). F-LEACH: a fuzzy-based data aggregation scheme for healthcare IoT systems. The Journal of Supercomputing, 78(1), 1030-1047.
18. Wang, X., Zhou, Q., & Tong, J. (2019). V-matrix-based scalable data aggregation scheme in WSN. IEEE Access, 7, 56081-56094.
19. Xue, K., Zhu, B., Yang, Q., Wei, D. S., &Guizani, M. (2019). An efficient and robust data aggregation scheme without a trusted authority for smart grid. IEEE Internet of Things Journal, 7(3), 1949-1959.
20. Zeng, Z., Liu, Y., & Chang, L. (2022). A robust and optional privacy data aggregation scheme for fog-enhanced IoT network. IEEE Systems Journal, 17(1), 1110-1120.

21. Zhang, J., & Dong, C. (2023). Privacy-preserving data aggregation scheme against deletion and tampering attacks from aggregators. Journal of King Saud University-Computer and Information Sciences, 35(4), 100-111.
22. Zhang, X., You, L., & Hu, G. (2022). An efficient and robust multidimensional data aggregation scheme for smart grid based on blockchain. IEEE Transactions on Network and Service Management, 19(4), 3949-3959.
23. Zhang, J., Zhang, W., Wei, X., & Liu, H. (2024). EPri-MDAS: An efficient privacy-preserving multiple data aggregation scheme without trusted authority for fog-based smart grid. High-Confidence Computing, 100226.
24. Hou, N., Xia, X., Wang, Y., & Zheng, Y. (2024). One shot for all: Quick and accurate data aggregation for LPWANs. IEEE/ACM Transactions on Networking.