Hybrid Fuzzy Data Aggregation and Optimization-Based Routing for Energy Efficiency in Heterogeneous Wireless Sensor Networks

Asaad A. Alhijaj[1], Ali K. Marzook[2], Dheyaa Mezaal Hussein[3], Jawad Alkenani[4*]
[1] Department of Computer Information Systems, College of Computer Science and Information Technology, University of Basrah, Basrah, Iraq
[2] University of Basrah- College of Engineering- Department of Electrical Engineering
[3] University of Technology Office of the President of the University, Baghdad, Iraq
[4] Department of Computer Science, Shatt Al-Arab University College, Basra, Iraq
Email: asaad.abdulhassan@uobasrah.edu.iq, ali.marzook@uobasrah.edu.iq, 11745@uotechnology.edu.iq, Jawadalkenani@sa-uc.edu.iq
[*]Corresponding author

Keywords FSMORP, network lifetime, routing, FDA, wireless sensor networks

Received: July 23, 2024

Wireless sensor networks (WSNs) are networks with many sensor nodes that are utilized for various purposes, including the military and medical. In hazardous circumstances, precise data aggregation and routing are essential, and the energy consumption of the sensors needs to be closely controlled. Nonetheless, there is a significant chance of redundant data because of external factors and nearby sensors. A multitude of information can be found in large datasets, some of it unnecessary and others useful. This redundancy negatively impacts performance in terms of redundant transmission and computing costs. However, data aggregation might help a network get rid of unnecessary data. In this work, we present a hybrid protocol called fuzzy data aggregation with fuzzy spider monkey optimization routing protocol (FDA-FSMORP) that represents an intelligent approach to collecting sensor data in HWSNs considering energy consumption. The results indicated that the suggested method beat in minimizing data latency our approach reduced energy consumption by 73% using energy more effectively when compared to our simulated outcomes.

Povzetek: Predstavljen je hibridni protokol FDA-FSMORP, ki z uporabo mehke optimizacije poveča energetsko učinkovitost heterogenih brezžičnih senzorskih omrežij in podaljšuje življenjsko dobo omrežja.

1 Introduction

Wireless Sensor Networks (WSN) are composed of numerous networked nodes, each of which can recognize and communicate changes in their immediate surroundings. Wireless sensor networks (WSN) have numerous potential applications, including smart buildings, the internet ecosystem, battlefields, industry, healthcare, and agriculture [1]. The longevity of the network decreases as sensors lose power. Overcoming these obstacles requires making the most efficient use of energy. Repetitive information is produced by nodes that are close to one another or receive input simultaneously [2], [3]. Consequently, a network's life energy is less depleted during data processing, transmission, and reception. Before being transferred to the sink via routing protocols, data is first collected and then aggregated using functions like sum, average, etc. [4]. This gets rid of the necessity to send the sink a single sensed value at a time. There are several methods for lowering the amount of data in WSNs, including compression at the cluster head (CH) or the sink. Alternatively, several mobile sinks may be used to aggregate data in a heterogeneous WSN using a clustering approach. This clustering technique is especially advantageous for assisting CH in identifying the N-sensors that comprise its cluster and in recognizing their CH on the N-sensors. Following the arrangement of the sensors proposes a smart approach to aggregate the sensing data in HWSNs to consider the energy consumption. After that, the work proposes a new routing protocol for HWSNs to send the aggregate data from the sensor to the sink through the CHs is used to figure out how to make sure the network lasts as long as possible, and that energy isn't wasted. To reason about the best way to route HWSNs both between clusters and within clusters, it looks at three routing metrics for each node. These three metrics are (the highest remaining energy within the node, the smallest number of hops, and the least amount of traffic within the node). Although rapid advancements were experienced, small sensors could perform tasks at higher levels, such as multimedia data processing and transmission [2]. Many WSN researchers have focused on energy-saving solutions, such as energy-aware aggregate and routing algorithms, which aim to reduce processing and communication resource requirements by limiting the use of WSNs to simple data-gathering and reporting applications.

The relevant works' approach, performance, and outcomes are compiled in Table 1. Therefore, choosing

Table 1: Summarization table on the related works.

Ref	Methodology	Performance/Results
[16]	• Two-tiered heterogeneous sensor network	• The authors examined a two-tiered heterogeneous sensor network in which low-power sensors are controlled by a high-power head sensor within clusters of sensors. To collect data and conserve energy, they polled sensors rather than allowing them to broadcast data at random. Additionally, they extended the network's lifespan by implementing two novel strategies: employing several frequency channels and sectorizing clusters.
[26]	• Fuzzy Logic	• The suggested method was assessed and contrasted with Bayesian classification, probabilistic neural networks (PNNs), and network performance metrics including packet loss, delay, jitter, and performance assessment. While the fuzzy technique generally performed better than the PNN and Bayesian approaches, all approaches properly identified the QoS categories. By accurately calculating the QoS of the network, one can have a better understanding of its performance.
[27]	• KF-SVM	• A multitude of information can be found in large datasets, some of it unnecessary and others useful. Performance is negatively impacted by this redundancy in terms of redundant transmission and computing costs. On the other side, redundant data in a network may be removed by data aggregation. This research introduces a novel technique called Kalman filter with Support vector machine (KF-SVM) to categorize, aggregate data, and eliminate noise in wireless sensor networks (WSNs). This improves network lifetime and efficiency.
[5]	• clustering technology	• Provides an energy-efficient routing protocol that reduces cluster head energy consumption by using a double cluster head approach and uneven clustering technology to address hot spot issues. Furthermore, a hybrid time-driven and energy-driven cluster head rotation technique is provided to balance the energy consumption between cluster heads and cluster members.
[20]	• UWSNs	• To quickly detect and report oil traces to the sink, the authors propose a novel routing protocol for the ocean floor that combines sleep-scheduling routing with two-dimensional UWSNs.
[21]	• K-NN	• The authors of propose a new routing strategy that may drastically reduce latency and power consumption throughout the whole network by fusing the K-NN algorithm with the clustering method. This proposal demonstrates how to use node classifications and the shortest distances between them to form clusters.
[22]	• Balanced Routing Protocol	• The creators of devised a revolutionary balanced routing technique that makes use of two distinct channels in order to reduce the strain on battery life. This idea gives each node two shortest paths to the sink in order to reduce network congestion.
[11]	• clustering HWSNs	• The authors presented a novel approach to clustering HWSNs that featured a useful technique for determining the cluster's head node, degree of sensor nodes, and remaining energy. The chaining method is also used for gathering and sending the information package.
[23]	• Grey Wolf Optimizer	• The offers a ground-breaking, energy-saving method built on an improved Grey Wolf Optimizer (GWO). In order to improve the optimum solution discovery in GWO, it takes into account a fitness value. This leads to a more balanced cluster structure and a more equitable distribution of CHs. Sensor node transmission distances are updated in accordance with the distances to the sink and CHs to reduce energy consumption.
[20]	• PSO	• The introduces a multi-hop routing protocol for uneven dynamic clustering (PUDCRP) based on PSO (Particle Swarm Optimization). As nodes fail in the PUDCRP protocol, the cluster distribution will change dynamically. Prospective CHs nodes can be located using the PSO approach.
[25]	• FDA-SMORP	Fuzzy Data Aggregation with Spider Monkey Optimization (FDA-SMORP), an energy-efficient method for data aggregation at the cluster head and routing to the sink, is described. By attempting to balance out each node's energy usage, these solutions aim to cause all of the nodes in a wireless network to run out of energy almost simultaneously. The proposed methods are compared with some of the most popular WSN systems to show their effectiveness in reducing route planning delay, balancing energy consumption, and prolonging network lifetime.

outside the network. In WSNs, the most challenging issue is increasing energy efficiency to extend the networks becomes difficult in the current cooperatively inside and lifetime [5]. While sensor networks are comparable in many ways to other distributed systems, they face a unique set of challenges and limitations. These constraints influence the design of a network, resulting in protocols and algorithms distinct from those of other distributed systems. The following subsections highlight some of the most significant WSN difficulties that have been resolved using some algorithms, Routing, and Communication Because data is transferred cooperatively within and outside the network, problems arise in selecting optimal and efficient communication paths. Many of the challenges of the WSN are captured succinctly in the context of transmission and orientation. These challenges include bandwidth, routing protocols, communication range, data rate, network topology, and packet loss rate. Numerous routing strategies have been devised to address the routing and communication difficulties of WSNs. WSNs may lower their power consumption by data aggregation, as described in [5], where a population-based method like the ant colony system enables researchers to organically traverse research space in optimization settings in search of the most valuable data. Each cluster leader receives a different seed vector from the sink node to account for the spread of the network. Through a series of intermediate nodes, clusters send their measurement data to the final sink node. In [13], he detailed an incremental approach to training support vector machines (SVMs) that attempted to filter out irrelevant data. Fisher's Discrimination Ratio may be used to differentiate between sets of aggregated data and sets of dispersed data (FDR). SVMs may be trained in less time since fewer data samples are required. Many moveable troughs were used to illustrate ways for efficient data aggregation, as shown in [14]. The statically sink-based method involves sending data packets through the network by dumping them over a series of intermediate nodes. That's why the fixed basin wastes so much power. The employment of a mobile sink to collect data reduces the network's energy consumption, extending its lifespan. Particle swarm optimization is used on active sensor nodes to minimize duplicate data, as recommended in [16] and [17], and incremental Naive Bayes Prediction may be used to reduce data in [15]. In [16] and [17], data aggregation is suggested by using compressive sensing technology. In [11], Fuzzy Dstar-Lite is a routing approach suggested by the authors to achieve optimal data routing in HWSNs. Moreover, it explains the UED issue in the network and highlights the need to go beyond the obstruction scenario. In [18], open mining is introduced as an effective and inexpensive data aggregation strategy. Many WSNs are used in this data mining process. One central node collects and transmits information from multiple peripheral nodes, in [5] With the help of a neural network comprised of self-organized maps, we can reduce data redundancy and remove anomalies. The process based on the density and similarity of the data is significantly simplified when cosine similarity is used in sensor node development. Researchers presented a clustering strategy for HWSNs in [19], [20], employing a novel method for selecting the cluster head nodes, the number of sensor nodes, and the remaining energy. The chaining approach is also used to gather and send information packages. They introduced the Spider Monkey Optimization Routing Protocol (SMORP), a swarm-based intelligence technique, in both the heterogeneous HWSNs [10] and the homogeneous WSNs [21]. Following a predetermined set of routing criteria, this technique determines the most efficient path through the network.

This paper is organized as follows: In Section 2, Research Importance, sensor networks are comparable in many ways to other distributed systems, they face a unique set of challenges and limitations. These constraints influence the design of a network, resulting in protocols and algorithms distinct. In Section 3, the organization of a heterogeneous network, there are two types of sensors (i.e. N-sensors and CH-sensors) used. Section 4 presents a hybrid proposal (FDA-FSMORP) that represents an intelligent approach to collecting sensor data in HWSNs considering energy consumption. Section 5 shows the simulation results of the proposed method. In Section 6, the paper's conclusion is finally stated.

2 Research Importance

In recent years, there has been a growing interest in the possibilities for sensor collaboration in data collection, processing, and sending to the sink. On the other hand, the resource-constrained nature of sensor nodes creates many challenges in developing, running, and maintaining sensor networks in the real world, which requires energy awareness at all layers of the networking protocol stack. Energy consumption is a crucial factor to consider when designing WSNs. So, a lightweight system is favored to conserve the sensor node's energy and preserve data protected from various vulnerabilities. Additionally, multiple obstacles complicate energy-efficient communication in WSNs. Therefore, using energy-efficient routing protocols, and balancing energy consumption overseer all the nodes are the most significant challenges to be overcome to extend the network lifetime.

1- Many services are hampered by the continuing difficulties in implementing networks. Many variables can be manipulated, and others cannot be manipulated. For this reason, properties affecting network deployment must be identified and controlled.

2- Data collection, communication quality, and data processing. Other things to think about are in terms of energy conservation, node state, and transmission method.

3- WSNs need a way to balance the power consumption on all nodes so that these nodes consume their entire energy and die at about the same time.

4- Next, it is required to maintain data confidentiality using the best lightweight block encryption with a routing protocol to achieve data confidentiality, integrity, and secure routing when transmitting sensor information from source nodes to the sink. While sensor networks are comparable in many ways to other distributed systems, they face a unique set of challenges and limitations. These constraints influence the design of a network, resulting in protocols and algorithms distinct from those of other distributed systems. The following subsections highlight some of the most significant WSN difficulties that have been resolved using some algorithms.

- **Energy consumption**

WSNs have some design challenges, including energy efficiency. Sensing, connectivity, and data processing are the three aspects of power usage that need improvement. The sensor node's lifetime is often influenced by battery capacity. When it comes to sensor networks, the most prevalent stumbling block is constrained energy budgets. Sensors are frequently powered by batteries, which must be replaced or recharged regularly. A rechargeable sensor node should be able to function until its job is completed or the battery must be replaced. The mission duration depends on the application.

- **Data aggregation**

Given that radio transmission is the primary energy consumption, one way to reduce this consumption is to reduce communication overhead via data aggregation. Rather than transmitting every individual node measurement to the sink, intermediary nodes consolidate the raw data into a manageable quantity of data packets containing meaningful information. By eliminating repetitive and unneeded data readings, data aggregation seeks to avoid redundant packet transfers and thereby reduce communication energy consumption.

- **Design and deployment**

In some applications of WSNs, the distribution of sensor nodes over the area of interest is done according to a plan, contrary to the random deployment where sensors are scattered by throwing them on the area (e.g., dropped from an airplane). The design process in WSNs has the aim of specifying the type, quantity, and placement of sensor nodes to be deployed in an environment to have complete knowledge of their functional status. Computational Intelligence techniques can be useful to handle the designing and planning of WSNs deployment.

3 Organization of HWSNs

Two types of sensors (N-sensors and CH-sensors) are utilized in the organization of a heterogeneous network. An HWSN requires many standard sensors (N-sensors) to be dispersed randomly throughout the area. Moreover, the network includes multiple sensor nodes with enough capacity to function as cluster heads (CH-sensors). The CH in this case needs to be deployed carefully, considering the computation of the distance between the sensor and the cluster head and the sink, to ensure that all N-sensors are safe and can be attached to at least one CH. In this study, clustering approaches that are employed in heterogeneous WSNs [7, 11] and homogeneous WSNs [22] are used.

Using their unique identifiers, the CH-sensors in this configuration send out broadcast signals that pinpoint their precise position. The CH-sensor whose unique identifier is the smallest in length will be ranked highest. After that, the N-sensors arrange the CH-sensors they've heard in order of loudness. Each N-sensor will give preference to the CH-sensor if it is a viable option. After that, the CH-sensor will begin selecting N-sensors for clustering. All clusters, no matter how big or little, are handled in the same way. The clustering process for HWSNs using this method is shown in Figure 1.

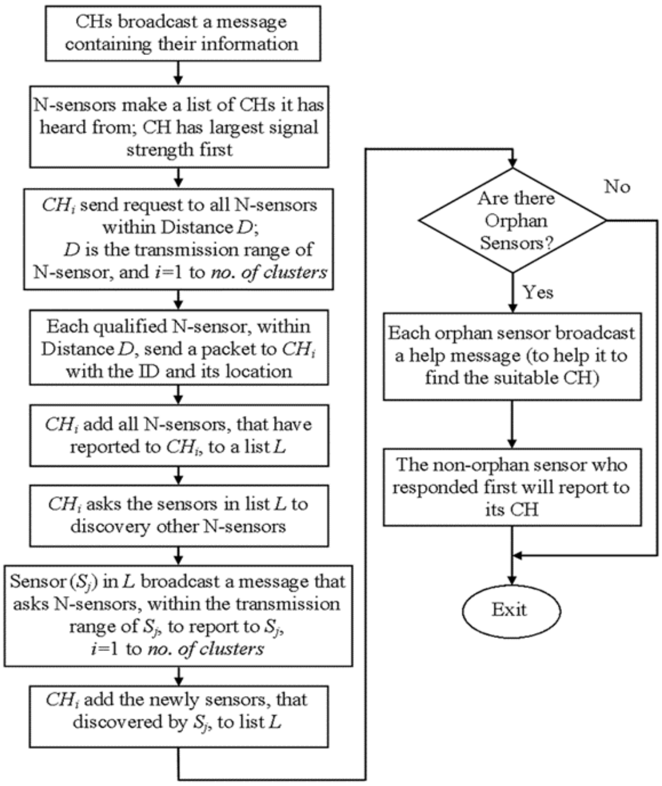

Figure 1: clustering to organize the nodes of a heterogeneous sensor [7].

According to WSNs research, several metrics or criteria can be used to find the routing path [13], [16], [19] .These metrics consist of the following Figure 2:

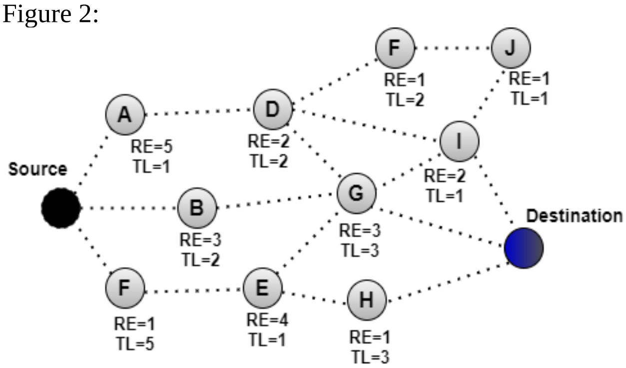

Figure 2: Path selection in WSNs using various routing parameters

- **Remaining energy (RE)**

In general, the most critical factor in the data routing of WSNs is energy efficiency. So, the remaining battery capacity of the sensor device is given significant consideration. Based on this metric, routing methods are designed to determine the path from the sender sensor to the destination by involving sensor devices with the highest remaining energy levels. Thus, sensors with more remaining energy would be utilized in the routing process more frequently than those with lower battery charge levels. A simple WSN with a small number of nodes is depicted in Figure 2-4, where a sender sensor (source) attempts to transmit a data packet to a sink node. The RE values within the nodes represent the remaining battery energy of the considered node. The optimal path would be (A-D-G) based on this concept, as it crosses the nodes with the highest remaining energy total (i.e., 19).

- **Minimum hop (MH)**

The minimal hops, also known as the smallest hop, is the most frequently employed measure in routing systems, where the routing method seeks to select the path that traverses the fewest forwarder sensors (hops) route to the sink. This criterion's central idea is that choosing the shortest path reduces end-to-end delay and energy depletion rates by incorporating the fewest relay nodes possible. The routing protocol would identify the path (B-G) in Figure 2-4, using this metric, which contains only two nodes and no shorter route.

- **Traffic load (TL)**

The amount of data traffics still pending in the queue of a sensor node is referred to as the "traffic load" or "intensity load" of that sensor. This traffic consists of both application-generated traffic and traffic sent from other sensors to be forwarded. When a sub-area experiences events more frequently than the entire deployment area, selecting the shortest transmission path for routing will cause implosion along the entire routing path. If a sensor node becomes overloaded with traffic, there is a high possibility of a queue overflow issue that results in the loss of crucial data. In addition, a high traffic load causes sensor nodes to lose energy quickly and excessively, resulting in a shorter network lifetime. The TL values within each node in Figure 2-4, represented the respective node's traffic load. Using this statistic, routing protocols would select a route (A-D-I) as the optimal path, as it traverses the nodes with the lowest total traffic load (i.e., 4).

To explore the issue of unbalanced energy utilization and to increase the network lifetime in WSNs, a new energy-efficient routing method called BAT Optimization Routing Protocol (BORP) is introduced in Chapter 4. The proposed routing method attempts to select an optimal path from the sender node to the sink, considering the above-mentioned routing metrics (RE, MH, and TL) and utilizing them in a balanced way to efficiently enhance the lifetime of WSNs.

4 The proposed approach

One of the biggest challenges facing a WSN is how to maintain the network lifetime. Considering this, the researchers presented several methods, including collecting data and how to direct it, considering the amount of energy consumed. In previous work, we have observed an improvement in wireless sensor networking using fuzzy spider monkey optimization routing protocol reducing costs and ensuring that power is distributed fairly, the FSMORP helps HWSNs last longer [24]. Then, in the second work, we observed the use of a hybrid protocol called fuzzy data aggregation with spider monkey optimization routing protocol (FDA-SMORP) [25]. In this work, we present a hybrid proposal (FDA-FSMORP) that represents an intelligent approach to collecting sensor data in HWSNs considering energy consumption. Next, the work proposes the FSMORP routing protocol for HWSNs to send the collected data from the sensor to the sink via cluster heads.

A. FUZZY DATA AGGREGATION (FDA) FOR HWSNS

The proposed method represents the process of aggregating data based on eliminating redundancy and extracting useful information. A similarity measure is a distance with dimensions that represent object attributes in the context of data mining. That example, if there is little separation between two data points, then the things they represent are quite similar to one another, and the converse is also true. One of the most frequent distance scales with which they are compared with the methodology presented (Jaccard similarity, Cosine similarity, Overlap Coefficient) is used by the vast majority of aggregation methods to evaluate the degree of dissimilarity between two items[23].

The suggested model's purpose is to guarantee that, when two sensors in proximity or sensing data at the same time submit an event, there is a high possibility that the same event will occur, resulting in an increase in data volume at the expense of network energy. In CHs, the FDA is used to efficiently aggregate data through redundancy removal and extract actionable intelligence.

The inference engine, which comprises a rule base and several strategies for inferring the rules (5^2=25 for the fuzzy rule base), deals with the fuzzy values. Similarity (n) is considered Medium if and only if FE (n) is high and SE (n) is low. A fuzzy inference engine processes all of these rules simultaneously. To extract a single, unambiguous value from the fuzzy solution space, the fuzzy values are first eliminated. This value represents the similarity value by which the data size is reduced. The CoG method of defogging is carried out by .Figure 3 shows a "fuzzy data aggregation" process into CHs**.**

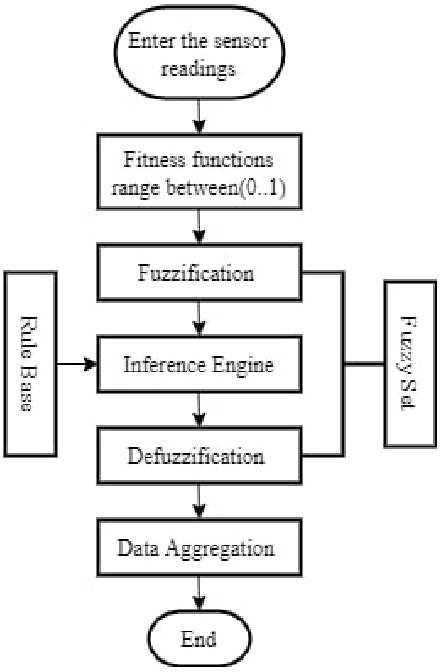

Figure 3: The fuzzy data aggregation process

Two events are evaluated for precision and similarity to evaluate the performance of the proposed method. There is a comparison of the fuzzy data aggregation technique to other approaches shown in Table 2.

Table 2: Accuracy similarity in the approaches for two sensor events

Algorithm/ Approach	Similarity	Accuracy
Jaccard similarity	0.25	0.48
Cosine similarity	0.87	0.83
Overlap Coefficient	0.88	0.84
Fuzzy Data Aggregation	0.96	0.93

The similarity criterion is chosen depending on the intended use of the data, and it is evident from this that a larger similarity percentage indicates more reliable data. In Figure 4, we can see the fuzzy data aggregating technique being used with a similarity criterion in the range (0,1).

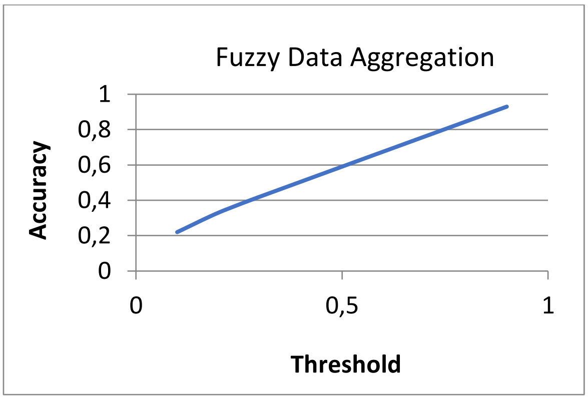

Figure 4: Percentage of accuracy after applying the threshold

B. FUZZY SPIDER MONKY OPTIMIZATION ROUTING PROTOCOL (FSMORP) FOR HWSNS

For HWSNs to last longer, the routing protocol is a critical issue. A breakdown in communication between the N-sensor and CH, or the CH and the sink, will occur if either sensor node loses power during the routing protocol. Consequently, there is often a deficiency in HWSNs throughout their useful lives. Since the lifetime of a HWSN is proportional to the power it receives from its sensors, the sensors must consume as little power as possible. By reducing costs and ensuring that power is distributed fairly, the FSMORP helps HWSNs last longer. The next FDMORP is by fusing the Fuzzy method with the Spider Monkey Optimization (SMO) [24] technique.

Based on the routing requirements, the FSMORP chooses the next hop to the sensor node (maximum remaining energy, fewest hops, and lowest traffic load). In this study, we assume that all N-sensors have the same range of transmission and start with the same amount of battery life. In (ii), every N-sensor knows its location and that of its CH and immediate surroundings. (iii). Each CH has the same battery-powered start-up time and transmission range (iv). Each CH knows its location as well as that of its neighbors (other CHs and the sink).

The routing schedule is created by the sink and broadcast to all connected nodes (N-sensors and CH-sensors). Each sensor uses an FSMORP to calculate the most direct path to the target node. There are two phases to the FSMORP:

1. SMO implementation in FSMORP

Here, a tree is evaluated using the SMO technique, and each node in the tree is given a fitness function value in the form of (N, Fit), where N is the collection of candidate nodes along the forwarding path and Fit is the set of fitness functions (n). It is the fitness function that determines which direction the tree node will go. As shown in Section 4.2, a fuzzy method is utilized to determine the value of the fitness function at each node.

In FSMORP, the generated routing path is utilized many times (rounds) before determining whether to continue using it based on the current state of each node along the route. Under these conditions, the sink would be aware of the real-time status of each node's battery life, location, and network use. To evaluate a neighboring node's viability, we apply Eq. (1). (ni).

$$fit(n_i) = fuzzy(RE(n_i), TL(n_i), D(n_i)) \qquad (1)$$

Node n's remaining energy (RE(n)), traffic load (TL(n), and distance to destination (D(n)) are represented as follows. The fitness value for node n may then be calculated using the fuzzy technique with these parameters as inputs. Then, using the data collected from the LLSM's neighbors, the GLSM selects the node with the highest probability P, where P is the probability value stated in Eq. (2):

$$P(n_i) = \frac{fit(n_i)}{\sum_{j=1}^{N} fit(n_j)} \qquad (2)$$

Where $P(n_i)$ is the probability associated with node n_i, fit (n_i) is the fitness associated with node n, and N is the number of neighbor nodes.

2. Fuzzy approach implementation in FSMORP

Here, we employ a fuzzy method to calculate the RE (n), TL (n), and D that together constitute the value of the fitness function at node n.

A single fit output parameter (n) is used in conjunction with three input parameters (RE (n), D (n), and TL (n), as seen in Figure 5) in the fuzzy technique. Discourse intervals of [0...0.5], [0...1], [0...10], and [0...1] are universal for all RE, D, TL, and fit, respectively.

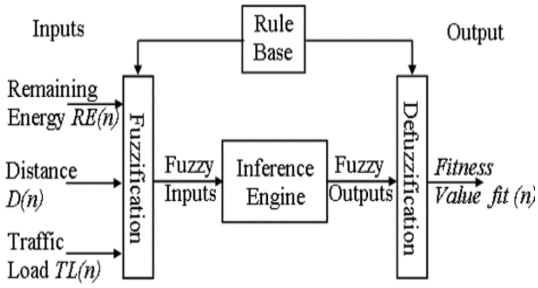

Figure 5: Fuzzy structure for FSMORP

Each input and output variable in FSMORP is represented by a set of five membership functions. As part of its data processing, FSMORP employs an inference engine, which is comprised of a rule base and several approaches for inferring the rules. In this document, we provide the IF-THEN rules that form the basis of FSMORP's fuzzy rule base, which includes a total of 5^3=125 rules.

All of these rules are run via a fuzzy inference engine concurrently. After being defuzzied, a solution's fuzzy space is collapsed into a single, clear output result. The value here represents the fitness of node s. Through the COG method [29] shown in Eq (3):

$$fit(n) = \frac{\sum_{k=1}^{N} U_k * c_k}{\sum_{k=1}^{N} U_k} \qquad (3)$$

In this expression, Uk represents the kth output of the rule set, and ck represents the center of the nth output membership function.

The FSMORP flowchart is seen in Figure 5, and it uses a hybrid SMO algorithm and Fuzzy technique to find the optimal routing route twice in inter-cluster (from N-sensor to its CH-sensor) and intra-cluster (from the CH-sensor to the sink) directions in a consistent, sequential fashion.

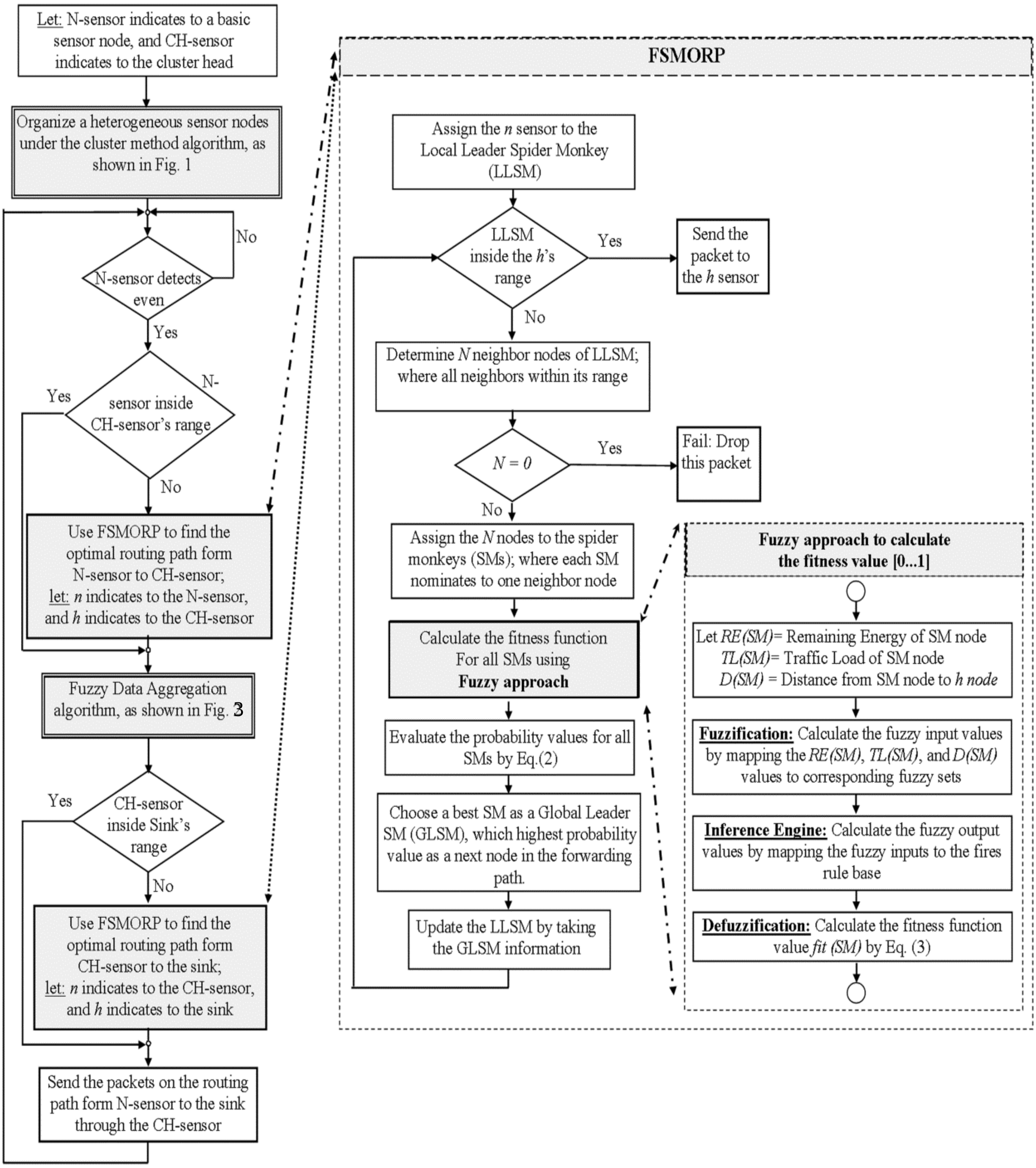

Figure 6: Flow-chart of the proposed method(FDA-FSMORP)

5 Performance evaluation

The life of the HWSNs can be extended by using the CH's fuzzy data aggregation method with the improved Routing Protocol to increase energy efficiency [10]. To see how well it worked, it was tested in three different approaches if the same routing metrics and the same environment were used in both.

1. Simulation Setup

Simulations are carried out in MATLAB R2010a (version 7.10) under Windows 7 (32 bits). The experiments are performed on a PC (ThinkPad T410i, China) with an Intel [R] Core [TM] i3 Processor running at 2.4 GHz and 2 GB of RAM is used to run the simulations. To make the network as realistic as possible, some parameters must be set in the system. Table 3 shows the Simulation parameters of the network, the network is the content of

1000 N-sensors, and 36 CHs randomly arranged within a 300 m x 300 m square topographical area.

The clustering technique groups N-sensors around CHs. Both systems use the [30] radio paradigm and have exhausted their transmission cycles (2000). Each approach produces a 2 KB packet length. On the other hand, all N-sensors and CHs start with the same starting energies of (0.5 J) and (2.5 J), with a sensed transmission of (20 m) and (80 m). The N-traffic sensor's load should be generated at random from 0 to 10. Every CH sensor has a [0..50] range.

Table 3: Simulation parameters

Parameter		Value
Area of topographical (meters)		300 m x 300 m
Location of the sink (meters)		(0, 150)
Length of control packets		2k
No. of transmission packets (rounds)		2 x 10^3
N-sensors	Number of nodes	1000
	Limit of transmission distance	20 m
	Initial energy	0.5 J
	E_{elec}	50 nJ/bit
	E_{amp}	100 pJ/bit/m^2
	Max. traffic in the node's queue	10
CHs	No. of nodes	36
	Limit of transmission distance	80 m
	Initial energy	2.5 J
	E_{elec}	100 nJ/bit
	E_{amp}	200 pJ/bit/m^2
	Max. traffic in the node's queue	50

2. Simulation Results

To obtain a more efficient sensor network, two approaches are proposed to reduce energy consumption; these two approaches are based on two scenarios. The first scenario represents the efficacy of the Enhanced Routing Protocol if a single sensor senses and transmits data at the same time as it is without aggregation. The proposed approach FSMORP was compared with SMORP. Thus, a balance in energy consumption was obtained.

The second scenario is a smart method to aggregate data to eliminate redundancy and conserve power consumption by assuming that two sensors sense and transmit data via the Enhanced Routing Protocol FSMORP. The robustness of the proposed FDA-FSMORP approach was compared with FSMORP and its effect on cluster heads was demonstrated.

- FSMORP

In this section, the number of sensors still functioning after each cycle of data transmission is used to compare the two systems' findings for network longevity. The proportion of N-sensors and CH-sensors that are operational in both the proposed system and the SMORP is shown in Figures 6 and 7, respectively. Therefore, the suggested solution surpasses the SMORP system in terms of how many nodes are still operational in the network. Here, and after 2000 packets have been sent via the network, the result of network lifetime attained with the suggested is about 15% higher than that of SMORP.

The network's lifespan is over as soon as any sensor (N-sensor or CH-sensor) in the network dies. The lifespan of the network has been extended thanks to the suggested approach. The time needed for the suggested method is more than that of the FDA-SMORP system.

The suggested method outperforms FDA-SMORP by a wide margin when it comes to minimizing energy consumption and maximizing network lifespan (see Figures 7 and 8).

When deciding between N-sensors and CH-sensors for routing, the fraction of unused power changes with the number of transmission cycles.

The suggested system outperforms the SMORP in terms of overall performance and efficiency thanks to the increase in the number of routes. Figure 7 and Figure 8 shows how the residual energy ratio for N-sensors and CH-sensors changes depending on the kind of transmission used. Keeping the network stable for as long as feasible, the suggested technique outperforms the FDA-SMORP.

The time lag that occurs when sending data packets, on the other hand, may make or break several uses. a comparison of the two systems with regard to the amount of time spent in simulation while still within the routing region. However, the suggested model seems somewhat bigger owing to the fuzzy logic processors when compared to the SMORP findings.

Figure 7 shows that the suggested system achieves low overall latency. Energy is conserved and data is sent more efficiently when there are fewer delays. That's why multipath routing is so important; it divides data packets up at the node level so that the network can function more smoothly and last longer even when traffic is heavy.

- FDA-FSMORP

In this section, the FDA-proposed algorithm is put into every cluster head. Thus, we notice the effect of the algorithm on clusters only, instead of the normal sensors. After each cycle of data transmission, the number of sensors still functioning in the network is counted to compare the network lifespan outcomes acquired using the two techniques. On this point, Figure 12 demonstrates the ratio of the CH-sensors, which are still alive in both the proposed system (FDA-FSMORP) and the FSMORP. Therefore, the FDA-FSMORP system performs better than the FSMORP system in terms of the number of active nodes in the network. In this case, after delivering 2,000 packets to two sensors through the network, the result of the network lifetime attained with the suggested is about 50% more than that of FDA-SMORP.

Whichever approach is used, the percentage of unused power in CH-sensors shifts as the number of transmission cycles increases. Overall, the suggested system's performance and efficiency are superior to those of the FDA-SMORP. shows how the residual energy ratio for

CH-sensors changes depending on the kind of transmission used. We can observe that the suggested solution outperforms the FDA-SMORP in terms of maintaining network stability for the longest feasible duration.

The network lifetime ends when any sensor (CH-or N-sensor) in the network dies. The network lifetime of the suggested system has increased. Table 4 presents a comparative analysis of the initial node failure in every system. The FSMORP system requires less time than the proposed system.

The suggested method outperforms the FSMORP technique in terms of balancing energy depletion and maximizing network longevity, as shown in Figures 7,8, and Table 4.

When we discuss the reasons behind any noticeable differences in performance, we note for example that the proposed method is more energy efficient, the reason is that the proposed method combines two clever techniques, the first is the FDA, which collects data at the head of the cluster based on the balance between the three metrics, where the role of fuzzy logic is to extract the best fitness value in terms of performance at the sensor based on the metrics (remaining energy, distance, traffic load). Thus, the data collection process is parallel throughout the network. The second method, FSMORP, determines the best path, considering that this method is also a hybrid (fuzzy logic with spider monkey algorithm). This method maintains a balance of energy consumption, and thus the lifetime of the sensor network lasts longer than in previous literature, due to better clustering and more efficient routing decisions.

Table 4: The number of rounds that begin with a dead node

Approaches	FSMORP	FDA-FSMORP
The lifetime of the first N-sensor that died (Rounds)	488	587
The lifetime of the first CH-sensor that died (Rounds)	698	733

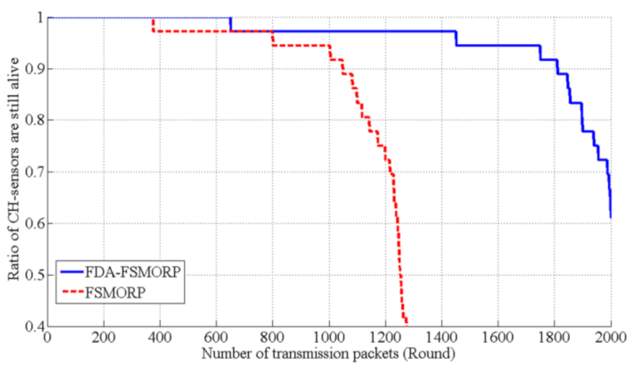

Figure 7: CH-sensors ratio remains alive (two N-sensors sent at a time)

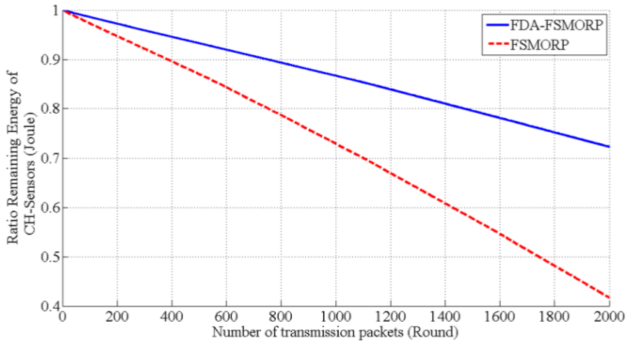

Figure 8: The energy ratio of the remaining CH-sensors (two N-sensors sent at a time)

6 Conclusions

Some of the data in large databases is helpful, while other data is utterly unnecessary. A data redundancy problem may arise if two sensors report the same event at the same time or within a little proximity to one another. The most effective use of energy is essential to overcoming these challenges. The information is compiled in the cluster header and then transmitted to the sink. On the other side, data aggregation has the potential to rid a network of unnecessary data, hence reducing data transmission while simultaneously increasing the network's lifespan. As a starting, this work suggests using a clever method called (FDA-FSMORP) which represents a hybrid approach to collect sensor data in HWSNs considering energy consumption. From what we can see from running simulations of the proposed model, FDA-FSMORP performs better than its competitors at both decreasing data latency and increasing network longevity.

References

[1] T. Kavitha and S. Silas, "A survey on energy harvesting routing protocol for WSN," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, no. Icoei, pp. 572–575, 2019, doi: 10.1109/ICOEI.2019.8862780.

[2] N. R. Roy and P. Chandra, "EEDAC-WSN : Energy Efficient Data Aggregation in Clustered WSN," *2019 Int. Conf. Autom. Comput. Technol. Manag.*, pp. 586–592, 2019, doi:10.1109/ICACTM.2019.8776679.

[3] J. Abdullah, M. K. Hussien, N. A. M. Alduais, M. I. Husni, and A. Jamil, "Data reduction algorithms based on computational intelligence for wireless sensor networks applications," *ISCAIE 2019 - 2019 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 166–171, 2019, doi: 10.1109/ISCAIE.2019.8743665.

[4] Hussein, Dheyaa Mezaal, et al. "Lifetime Maximization Using Grey Wolf Optimization Routing Protocol with Statistical Technique in WSNs." *Informatica* 47.5,2023,doi.org/10.31449/inf.v47i5.4601.

[5] I. Ullah and H. Y. Youn, "A novel data aggregation scheme based on the self-organized map for WSN," *J. Supercomput.*, vol. 75, no. 7, pp. 3975–3996, 2019, doi: 10.1007/s11227-018-2642-9.

[6] T. Dao, T. Nguyen, J. Pan, Y. U. Qiao, and Q. Lai, “Identification Failure Data for Cluster Heads Aggregation in WSN based on Improving Classification of SVM,” *IEEE Access*, vol. PP, p. 1, 2020, doi: 10.1109/ACCESS.2020.2983219.

[7] I. S. Alshawi, L. Yan, W. Pan, and B. Luo, "Fuzzy chessboard clustering and artificial bee colony routing method for energy-efficient heterogeneous wireless sensor networks," International Journal of Communication Systems, vol. 27, no. 12, pp. 3581-3599, June 2014, doi: 10.1002/dac.2560.

[8] W. K. Yun and S. J. Yoo, “Q-Learning-based data-aggregation-aware energy-efficient routing protocol for wireless sensor networks,” *IEEE Access*, vol. 9, pp. 10737–10750, 2021, doi: 10.1109/ACCESS.2021.3051360.

[9] N. Chandnani and C. N. Khairnar, “Efficient Data Aggregation and Routing Algorithm for IoT Wireless Sensor Networks,” *IFIP Int. Conf. Wirel. Opt. Commun. Networks, WOCN*, vol. 2019-Decem, 2019, doi 10.1109/WOCN45266.2019.8995074.

[10] Alkenani, Jawad, and Khulood Ahmed Nassar. "Network Monitoring Measurements for Quality of Service: A Review." *Iraqi Journal for Electrical & Electronic Engineering* 18.2,2022, doi: 10.37917/ijeee.18.2.5.

[11] I. S. Alshawi, A.-K. Y. Abdulla, and A. A. Alhijaj, “Fuzzy star-lite routing method for energy-efficient heterogeneous wireless sensor networks,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 19, no. 2, pp. 1000–1010, August 2020, doi: 10.11591/ijeecs.v19.i2.pp906-916.

[12] R. Misra and C. Mandal, “Ant-aggregation: ant colony algorithm for optimal data aggregation in wireless sensor networks,” in *2006 IFIP International Conference on Wireless and Optical Communications Networks*, pp.5-pp, 2006, doi: 10.1109/WOCN.2006.1666600.

[13] X. Shen, Z. Li, Z. Jiang, and Y. Zhan, “Distributed SVM classification with redundant data removing,” in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical, and Social Computing*, pp. 866–870, 2013, doi: 10.1109/GreenCom-iThings-CPSCom.2013.152.

[14] A. Muthu Krishnan and P. Ganesh Kumar, “An Effective Clustering Approach with Data Aggregation Using Multiple Mobile Sinks for Heterogeneous WSN,” *Wirel. Pers. Commun.*, vol. 90, no. 2, pp. 423–434, 2016, doi: 10.1007/s11277-015-2998-6.

[15] P. D. Ganjewar, S. Barani, and S. J. Wagh, “Data reduction using incremental Naive Bayes Prediction (INBP) in WSN,” *Proc. - IEEE Int. Conf. Inf. Process. ICIP 2015*, pp. 398–403, 2016, doi 10.1109/INFOP.2015.7489415.

[16] D. gan Zhang, T. Zhang, J. Zhang, Y. Dong, and X. dan Zhang, “A kind of effective data aggregating method based on compressive sensing for wireless sensor network,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2018, no. 1, 2018, doi: 10.1186/s13638-018-1176-4.

[17] M. I. Adawy, S. A. Nor, and M. Mahmuddin, “Data redundancy reduction in wireless sensor network,” *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 1–11, pp. 1–6, 2018, doi: 10.1002/dac.2560..

[18] H. Ramezanifar, M. Ghazvini, and M. Shojaei, “A new data aggregation approach for WSNs based on open pits mining,” *Wirel. Networks*, vol. 27, no. 1, pp. 41–53, 2021, doi: 10.1007/s11276-020-02442-9.

[19] M. Pandey, L. K. Vishwakarma, and A. Bhagat, “An energy efficient clustering algorithm for increasing lifespan of heterogeneous wireless sensor networks,” in *International Conference on Next Generation Computing Technologies*, pp. 263–277, 2017, doi: 10.1007/978-981-10-

[20] M. Al Mazaideh and J. Levendovszky, “A multi-hop routing algorithm for WSNs based on compressive sensing and multiple objective genetic algorithm,”*J. Commun. Networks*,no.99, pp. 1–10, 2021,doi: 10.23919/JCN.2021.000003.

[21] Jabbar, Ali H., and Imad S. Alshawi. "Spider monkey optimization routing protocol for wireless sensor networks." International Journal of Electrical & Computer Engineering (2088-8708), Vol. 11, No. 3, pp. 2432~2442, June. 2021, doi: 10.11591/ijece.v11i3.pp2432-2442.

[22] O. Younis and S. Fahmy, “HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks,” *IEEE Trans. Mob. Comput.*, vol. 3, no. 4, pp. 366–379, October. 2004, doi: 10.1109/TMC.2004.41.

[23] G. Sahar, K. A. Bakar, F. T. Zuhra, S. Rahim, T. Bibi, and S. H. H. Madni, “Data Redundancy Reduction for Energy-Efficiency in Wireless Sensor Networks: A Comprehensive Review,” *IEEE Access*, 2021, doi:

[24] Zaboon, Khalid Hameed, Nagham Mumtaz Kudhair, and Imad S. Alshawi. "Fuzzy spider monkey optimization routing protocol to balance energy consumption in heterogeneous wireless sensor networks." *Indonesian Journal of Electrical Engineering and Computer Science* 29.2 (2023): 921-930.2." *Informatica* 46.7,2022,doi: 10.11591/ijeecs.v29.i2.pp921-930.

[26] Altmemi, Dhuha Kh, Abdulmalik Adil Abdulzahra, and Imad S. Alshawi. "A new approach based on intelligent method to classify quality of service." *Informatica* 46.9,2022,doi.org/10.31449/inf.v46i9.4323.

[27] Kheerallah, Yousif Abdulwahab, and Jawad Alkenani. "A new method based on machine learning to increase efficiency in wireless sensor networks." *Informatica* 46.9,2023,doi.org/10.31449/inf.v46i9.4396.