SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I3P101

Volume 12 Issue 3, 1-11, March 2025

Original Article

Modelling a Novel Random Trusted Multi-Aggregation Algorithm (RTMA^2) for Enhances Data Aggregation

R. Nandha Kumar[1], P. Srimanchari[2]

[1,2]*Department of Computer Science, Erode Arts and Science College, Erode, Tamilnadu, India.*

[1]*Corresponding Author : prof.nk.07@gmail.com*

Received: 02 January 2025 Revised: 03 February 2025 Accepted: 01 March 2025 Published: 29 March 2025

Abstract - *It is essential to concentrate on the data aggregation process in order to improve the security of Wireless Sensor Networks (WSNs) for safe data transfer at both the Cluster Head (CH) and Base Station (BS). During data transmission, this technique combines data at each router. However, relying on energy-wasting nodes for this function can significantly shorten the network's lifespan. Therefore, improving the data aggregation in WSNs is essential to increasing their energy efficiency. This refinement can be made at the CH and the BS using various mathematical techniques. Sensors are vulnerable to adversaries' attacks when positioned over wide areas. These are the primary attackers who can introduce erroneous data into susceptible sensors during the data aggregation phase, resulting in inaccurate judgments at particular base stations. When it comes to such attacks, traditional methods of data averaging are unreliable. Observing how nodes behave during each data aggregation round is crucial to eliminate the inaccurate data introduced by these malicious parties precisely. This approach reduces the effect that inaccurate data has on the results. The Random Trusted Multi-Aggregation Algorithm (RTMA^2) improves data aggregation by maximizing the energy efficiency of clustering by fusing a safe data aggregation technique with the capability to confirm data dependability. However, the entire data aggregation process is at risk since the aggregator has been compromised. Therefore, an aggregation protocol that maintains energy efficiency and security while also being resistant to compromised aggregators must be developed. In addition to reducing the dangers associated with compromised aggregators, The novel trust-based multi-aggregation protocol (RTMA^2) shows how, in comparison to current methods, RTMA^2 improves the energy efficiency, security, and adaptability of WSNs.*

Keywords - *Sensor networks, Data aggregation, Trusted model, Random approach, Energy efficiency.*

1. Introduction

Data aggregation is a key strategy for improving how much energy is used in networks, as it reduces the need for sending data frequently, making the network last longer [1]. Data aggregation is the process of collecting and summarizing data from multiple sources or datasets into a unified format for analysis or reporting. The main objective of this method is to maintain equilibrium in energy consumption, which is necessary to extend the network's lifespan. The process of using less energy to do the same work or generate the same result is known as energy efficiency. It entails maximizing energy use by eliminating waste and enhancing the functionality of energy-related appliances, systems, and procedures. Unbalanced energy utilization can jeopardize the stability and security of the network [2]. Increasing the connection between WSN sensors and receiving points can assist in postponing the network's division into smaller segments by ensuring that energy is distributed evenly among the sensors. WSNs comprise numerous tiny sensors intended to monitor various environmental factors in real-time [3]. One of the main challenges in WSNs is ensuring that these sensors can collect all the data they require while consuming the least amount of energy possible because of their restricted energy. Recent advances in Compressive Sensing (CS) technology have addressed the issue of excessive duplicate data and the need to regulate the number of data points in WSNs in order to save energy [4]. With the use of this innovative technology, signals that are likely to be accurate can be collected and assembled using fewer samples than the Nyquist sampling theorem indicates.

Compressive sensing enables the main signal to be simplified at the nodes by utilizing the signal's inherent sparsity. With this method, there is no longer a need to collect and compile data at the receiving point [5]. Compressive sensing has demonstrated encouraging outcomes in various fields, including image processing. Sensors closer to the central receiving point must send out a lot more packets than those at the network's outskirts when compressive sensing is not used for data collection. Compressive sensing allows each node to send out only a few packets for data collection [6]. As a result, the total amount of data transmitted via the network is precisely proportional to the square of the number of nodes (MN). However, the overall volume of data is still very significant, even with fewer packets being transmitted. A mixed strategy has been proposed by studies, in which nodes nearer the network's endpoints communicate their data without employing compressive sensing, while nodes closer to the primary receiving point use compressive sensing for data transfers. The least energy data aggregation tree was developed as a result of improving data collection through the refinement of this mixed technique [7].

Prior studies have effectively integrated compressive sensing concepts into data transmission line design. Data grouping techniques are particularly preferred for implementing compressive sensing in grouped networks because they provide numerous advantages over conventional routing techniques [8]. These grouping algorithms typically result in an even task distribution when compared to conventional data collection techniques. Additionally, previous studies have not adequately examined how to manage location data and optimize node placement [9], two aspects that might significantly lower the energy consumption of WSNs for various IoT. Studies that have examined the Toeplitz matrix's properties have discovered that it satisfies a number of particular requirements. The sparse matrices employed in the compressive sensing framework are inherently sparse due to the high degree of similarity among the data points within each cluster. By reducing the number of independent random variables, this intrinsic sparsity improves the computing speed and efficiency of the compressive sensing process [10].

The most recent set of research has examined a variety of sparse random matrix structures, including quasi-Toeplitz matrices, Toeplitz matrices, semi-Hadamard matrices, and chaos-Toeplitz matrices, successfully demonstrating that they satisfy the requirements of the Restricted Isometry Property [11]. Following these foundational studies, scholars have expanded on this concept by developing a variety of matrices that satisfy the RIP condition. In response to these developments, further research [12] has concentrated on improving Toeplitz matrices by applying Singular Value Decomposition (SVD) [13], especially in the context of WSNs [15].

These are unique features of the 4s research is its focus on the cluster structure present in WSNs. It initiates a cutting-edge data aggregation approach that uses sparse compressive sensing. This technique divides the data aggregation process into two steps: within the clusters, those sink node selects a relevant seed vector that reflects the network's distribution and sends this to each CH. Then, the CH creates a customized matrix based on the received main vector, which is used for data collection with compressive sensing technology. This method allows the collection of measurement features from the clusters to the central sink node through a pre-defined multi-hop routing tree. Performance evaluations and comparative analyses of the experimental outcomes against other methods demonstrate that this approach is efficient and superior in reducing the system's overall energy consumption, thereby increasing its overall durability arrays effectively demonstrating their compliance with the clustering standard. Based on these results, the researchers have broadened the discussion by suggesting different arrays that meet the aggregation requirement. Building on these observations, this work has developed methods to improve the aggregation model through matrix vectors, especially in WSN environments.

The distinctive feature of this research is its focus on the cluster structure of WSNs. It introduces an innovative method for data collection that uses sparse compressive sensing. This method divides the data collection into two phases: within the clusters, the sink node selects an appropriate seed vector considering the network's distribution and passes this information to each cluster head. The challenges posed by the limited capabilities of SNs, especially in large networks where many nodes operate as data relays, are significant. Data broadcasting through wireless networks consumes more energy in sensor systems, so it's crucial to manage energy use efficiently. To address this issue, several approaches have been proposed aimed at optimizing energy use and enhancing data integration methods.

Many current algorithms are designed with a focus on minimizing energy consumption, but they overlook adaptive mechanisms that can respond dynamically to changing network conditions, sensor node failures, and varying traffic loads. These algorithms are often static, which reduces their effectiveness in real-world, large-scale, or dynamic environments. Furthermore, the majority of these protocols do not offer solutions for balancing the trade-off between energy efficiency and security, especially when dealing with compromised nodes or malicious attacks. This limits their applicability in mission-critical applications where data integrity and security are paramount. To address these challenges, this paper introduces the RTMA^2 protocol, which is designed to enhance data aggregation performance in WSNs by combining adaptive algorithms with trust-based security mechanisms.

RTMA^2 introduces a multi-agent system that dynamically monitors the behavior of sensor nodes and aggregators, assigns trust scores based on their reliability, and adjusts aggregation strategies to prevent data tampering and resource depletion. By integrating a trust-based model with adaptive algorithms, RTMA^2 offers a secure, energy-efficient, and scalable solution for data aggregation in WSNs.The proposed RTMA^2 protocol fills the research gap by addressing the vulnerabilities of traditional data aggregation algorithms, particularly in mitigating risks associated with compromised aggregators. RTMA^2 not only reduces the energy consumption of WSNs but also incorporates robust security measures to ensure data integrity and prevent malicious activities. Additionally, the adaptive nature of RTMA² allows it to dynamically adjust aggregation strategies in response to changing network conditions, enhancing its performance in large-scale and dynamic environments. The work aims to:

- ✓ Detect the limitations of existing data aggregation algorithms like security, energy efficiency, and scalability.
- ✓ Develop a novel trust-based multi-agent aggregation protocol (RTMA²) to mitigate risks associated with compromised aggregators.
- ✓ Demonstrate how RTMA² enhances the energy efficiency, security, and adaptability of WSNs compared to existing solutions.

The rest of the research is organized as follows: Section 2 gives a wider evaluation of existing techniques, the methodology is described in Section 3 with numerical findings in Section 4 and the conclusion is explained in Section 5.

2. Related Works

Underground WSNs and Actuator Networks (WSANs) are two examples of the several types and classifications of WSNs. Wireless sensor networks, multimedia WSNs, mobile WSNs, underwater WSNs, and terrestrial WSNs are all efficient ways to enhance algorithm performance and lower energy consumption [16-18]. Long routes can be divided into smaller, easier-to-manage pieces to drastically reduce energy consumption, as the quantity of energy required in wireless networks is dependent on distance [19]. However, other issues like data integration, reliability, and load balancing should also be taken into account while building routes [20]. Node grouping is a popular method for improving route efficiency and addressing these problems. Numerous clustering techniques have been developed over time; some, such as LEACH and HEED, date back to the beginning of the 2000s [21-23]. These strategies seek to improve important network functions like load balancing while simultaneously reducing energy consumption.

This section explores the other clustering methods used in WSN based on a detailed review paper. Unlike previous reviews, this work does not focus on the complex technical details of clustering methods, such as the algorithm's complexity or the specific techniques used. Instead, the author looks at the most impactful clustering techniques in WSN [24], assessing them based on their goals and their ability to improve network characteristics like node diversity and mobility. This shows that Table 1 shows you an overview of the main clustering techniques in WSNs [25], encompassing their objectives and the network functionalities they improve. This thorough review aids in selecting methods according to certain clustering criteria and objectives [26]. At the conclusion of each approach, clustering goals that are exclusive to that technique and not in the standard list are highlighted. This table, together with the objectives and network aspects they cover, is thought by the author to be the most comprehensive compilation of current clustering approaches in WSAN [27]. A number of variables, including energy sources, processing power, and network connectivity, affect the efficiency of the method. Data transport can be supported in resource-rich areas to enable nodes with limited resources to last longer on their own. However, it can be difficult to include nodes with low resources in group duties like data transfer [28]. Reduced Quality of Service (QoS), a decrease in energy supply, and the possibility of individual node failure are the causes of including new nodes in the network. A network's efficiency depends on the diversity of its components since it enables resources to be redistributed across various nodes in order to complete particular tasks. However, networks with multiple components have a difficult time efficiently allocating and managing these resources. Since those with greater resources are typically better suited for this crucial role, having a diversity of nodes can significantly improve the selection of CH nodes in the grouping context.

Finding the appropriate CH nodes, however, might also present new difficulties, which are covered in more detail in the sections that follow. In addition to facilitating hierarchical data fusion and compression, inter-cluster routing plays a critical role in caching and improved load balancing [29]. In order to facilitate communication between members of different clusters without requiring a Base Station (BS), inter-cluster broadcasting aims to build a network topology that links them. Instead of delivering data straight to the BS, this routing technique enables CHs to transit information indirectly through other CHs. There are a number of inter-cluster routing techniques that can be based on well-known protocols like AODV or integrated. Inter-cluster broadcasting enhances network connectivity by connecting clusters, which increases efficiency by facilitating resource sharing and the completion of distributed activities [30]. Intra-cluster routing is purposefully left out of the analysis because its goal is to minimize energy consumption, whereas inter-cluster routing concentrates on connecting clusters to connect all of the network's nodes. In terms of routing tactics, the author focuses on clustering methods that expand the number of connections between nodes, primarily to enhance energy efficiency via intra-cluster routing [31].

LEACH performs poorly in large-scale networks due to its reliance on direct communication from cluster heads to the BS, leading to increased energy utilization in distant nodes. Chain formation can cause significant delays in data transmission, especially in long chains. Like LEACH, HEED struggles in large-scale networks where long-range communication between CHs and the base station increases energy consumption. TEEN's focus on threshold-based reporting results in sparse data collection, which may be inadequate for continuous monitoring applications [31]. While the existing data aggregation algorithms have been successful in reducing energy consumption and improving data transmission efficiency, they have limitations in scalability, security, energy management, and adaptability. Addressing these limitations through adaptive, secure, and scalable approaches will be crucial in developing more efficient and robust aggregation protocols for modern WSNs.

3. Methodology

The framework for measurement can be broken down into various smaller frameworks, each targeting a particular group or cluster. The i^{th} the smaller framework is symbolized as ϕ^{H_i} with CH_i being the head of this framework and the data vector linked is referred to as x^{H_i}. CH_i is tasked with determining the measurement values for the data vector x^{H_i} using the specific framework's parameters. After CH_i has calculated the expected measurement values, this data is sent to the final node through the main tree structure that links the cluster leaders to the final node. For the purpose of this example, the author will consider that all nodes are divided into four categories, as groups of 5 to 8 nodes can be efficiently grouped into four categories. These categories are associated with a main aggregation tree. The data vector x can be given as $[x^{H_1}, .. x^{H_4}]^T$ and the measurement framework is symbolized as $[\phi^{H_1}, \phi^{H_2}, \phi^{H_3}, \phi^{H_4}]$. The premises presented in this study are accurate, suggesting that the conclusions are relevant to real-life situations, and these conclusions can be proven through real-world requests.

$$y = \phi x = [\phi^{H_1}\phi^{H_2}\phi^{H_3}\phi^{H_4}] = \begin{pmatrix} x^{H_1} \\ x^{H_2} \\ x^{H_3} \\ x^{H_4} \end{pmatrix}$$
$$= \sum_{i=1}^{4} \phi^{H_1}.\ x^{H_i} \quad (1)$$

As described in Equation (1), the expected coefficient of the measurement framework is obtained by aggregating all coefficients within each cluster. Thus, at each phase, the cluster leader generates estimated coefficients, which are relayed to the final node by each cluster leader. After receiving M rounds of estimated coefficients from the cluster leaders, the final node can reconstruct the original data. The compression ratio is represented as $\rho = M/N$ reflects the relationship between the measurement value M obtained through various compressive sensing techniques and the overall length N of the received signal. This ratio serves as an indicator of compression efficiency throughout the network. The relative reconstruction error is calculated as $\varepsilon = \frac{||d - d^{\Lambda}||^2}{||d||_2^2}$, which measures the ratio of the absolute error to the actual value. In this context, d denotes the true distance from $ith\ node$ to CH, while d^{Λ} indicates the measured distance from ith node to its CH.

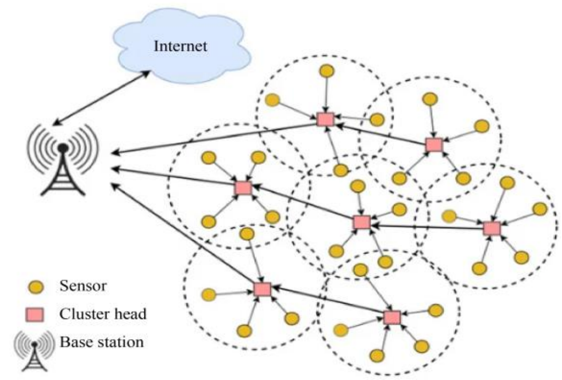

Fig. 1 Clustering in WSN

3.1. Sensing Network

In the sensing network, the application of the compressive sensing method can greatly diminish the energy utilization of each node; however, it is closely associated with the measured value M in the realm of sparse sensing. An increase in M results in increased energy usage by the nodes. To mitigate this challenge, this work proposes a novel strategy that integrates compressive sensing with data aggregation techniques. This strategy consists of four primary steps: first, partitioning the network into clusters; second, establishing a routing tree for inter-cluster communication; third, consolidating data within clusters using compressive sensing; and finally, transmitting data from cluster heads to the destination. The methodology for constructing the routing tree and enhancing compressive sensing within clusters is detailed below. The following assumptions regarding the network are made, which are generally applicable and facilitate the understanding that the results are relevant to real-world situations:

- There are N nodes randomly placed within a circular sensing area of radius L with the sink node located in the middle (refer to Figure 1).
- The sink node has sufficient storage capability to execute data.
- The primary energy levels and data broadcasting speeds of each sensor node are uniform across the network.
- Each node has technology that enables it to recognize its position relative to other nodes.

Theorem 1: Within WSN, the components are spread out randomly; the act of gathering data among clusters makes use of sparse matrices. This becomes especially efficient when the cluster's central component is placed close to all other cluster nodes, reducing the energy needed for data collection. To illustrate, imagine a cluster made up of m_j' components with a measurement matrix with a low density (s) in the compressing sensing method.

$$m_j' = \sum_{i=1}^{m_j} s * 1 = m_j s \quad (2)$$

During each data collection phase, the cluster typically sees m'_j components participate in data aggregation. It's clear that only m'_j components have to send their data signals at these times, leading to the cluster's central node receiving m'_j signals. Therefore, the total energy required for data collection in the j^{th} cluster can be represented as:

$$\begin{aligned}\bar{E}^j_{intra} &= \sum_{i=1}^{m'_j} E^i_{Tx}\big(k, \quad E(d_i)\big) + m'_j E_{Rx}(k) \\ &= k\sum_{i=1}^{m'_j}\left(E_{ele} + \varepsilon_{amp}(d_i^2)\right) + m'_j\, kE_{ele} \\ &= 2m'_j\, kE_{ele} + k\varepsilon_{amp}\sum_{i=1}^{m'_j} E(d_i^2)\end{aligned} \tag{3}$$

Where, $E^i_{Tx}(k, E(d_i))$ shows the amount of energy needed by the i^{th} component to send k bits of data to the CH and $E(d_i)$ represents the expected distance from i^{th} component to CH. This equation demonstrates that the average amount of energy used depends on $E(d_i^2)$.

$$f(x,y) = \begin{cases}\dfrac{1}{b^2} & x \in \left(-\dfrac{b}{2}, \dfrac{b}{2}\right), y \in \left(-\dfrac{b}{2}, \dfrac{b}{2}\right) \\ 0 & else\end{cases} \tag{4}$$

$$\begin{aligned}E(d_i^2) &= E((x-x_0)^2 + (y-y_0)^2) \\ &= \int_{-b/2}^{b/2}\int_{-b/2}^{b/2} \frac{1}{b^2}\big((x-x_0)\big)^2 \\ &\quad + (y-y_0)^2)\, dxdy\end{aligned} \tag{5}$$

$$= \frac{b^2}{6} + (x_0^2 + y_0^2) \geq \frac{b^2}{6} \tag{6}$$

Consider a cluster arranged in a square shape, with a side length labelled as b and its center at coordinates ($x_0 = y_0 = 0$). The likelihood of a component being close enough to the cluster's central node can be explained by the function$f(x,y)$. This equation is valid only if $x_0 = y_0 = 0$, indicating the cluster's center is exactly in the middle of the cluster. When examining the network as split into N_C clusters, each network cluster has a designated central node, while the others are linked to the closest cluster head. It's also suggested that each component can adjust its energy levels based on the distance it needs to communicate. As a result, the energy used by the component n_i to communicate with the component n_j is given by $P_{ij} = d_{ij}^{\propto}$. The variable α varies with the channel's characteristics and usually ranges from 2 to 4. For the discussion at hand, the methoduses $\alpha = 2$, which is a common value in traditional WSN setups. By the end, the work use the regulated rebuilding error as a measure for evaluating the quality of the Compressed Sensing (CS) signal.

3.2. Cluster Routing Model

The data transmission between clusters involves sending information through a series of intermediate points from one cluster's central node to another (NoH), where each node calculates its own worth based on communication range and another layout in cluster central nodes across the network. Figure 2 states if cluster central nodes pass on measurement data along the shortest path connecting other cluster central nodes, the energy used for data transfer between clusters will be reduced. Whenever a cluster central node collects measurement data, it will receive $h - 1$ data packets, and the energy spent on data transfer between clusters is defined as:

$$\begin{aligned}E_{inter} &= \sum_{i=1}^{h} E^i_{tx}(k, d_i) + (h-1)E_{Rx}(k) \\ &= k\sum_{i=1}^{h}\left(E_{ele} + \varepsilon_{amp} d_i^2\right) + (h-1)\, kE_{ele} \\ &= (2h-1)\, kE_{ele} + k\varepsilon_{amp}\sum_{i=1}^{h} d_i^2\end{aligned} \tag{7}$$

In this scenario, let d_i represent the space of the i^{th} information packet's transmission. The method mentioned earlier depicts that when h & k are constant, the overall energy used is determined by adding the squares of these distances, represented as $\sum_{i=1}^{h} d_i^2$. This work proposes a step-by-step method to create a distributed routing system for clusters. It's assumed that all cluster central nodes have a similar broadcasting range, which is called$R$. Within this range, these nodes can communicate with each other. Each cluster central node will share information about the number of intermediate points (NoH) from itself to the final destination node and its adjacent nodes. To start, the NoH for the cluster central node that includes the final destination within its communication range is set at 1. In the next step, these cluster central nodes will pass on their NoH to their adjacent nodes, setting a NoH of 2 for those that don't yet have a NoH. This procedure is executed iteratively until all cluster central nodes are assigned to the routing paths. The algorithm is given as follows:

Algorithm 1:
1. While (route no variation in CH routes)
2. Node (sink) = 0, CH, $i \in N_c$;
3. Initialize the neighborhoodnode-set based on CH;
4. if node distances < R; $j \in R$
5. when neighbourhood node (j) = min {node intersection of neighbourhood};
6. CH routes with j^{th} CH;
7. Add new node to neighborhood nodes;
8. End if
9. End while

3.3. Data Aggregation with Random Trusted Multi-Aggregation Algorithm (RTMA²)

Once the routing tree is set up across different clusters, this work employs compressive sensing technology to make data gathering more efficient within these clusters. The strong relationship between data points in a cluster enables us to significantly cut down on the number of measurements by using a chance sparse matrix. In the conventional approach to data collection with sparse sensing, the matrix needed for this technique is initially provided by the cluster leader. After the data collection is complete, the cluster leader is tasked with sending both the data and the matrix to the final destination. However, since the arbitrary sparse row can be explicitly developed by the final destination using a seed vector of the sparse sensing technique, each cluster leader can create a unique sub-matrix by utilizing the seed vector from the final destination. The process is carried out in the next steps:

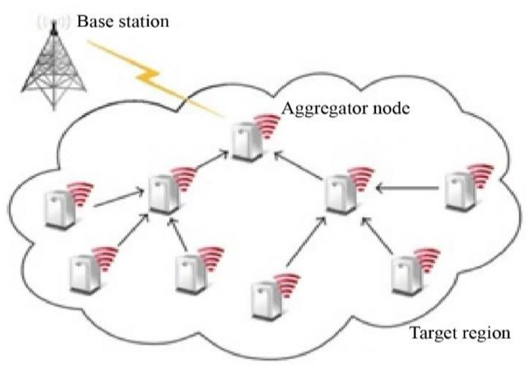

Fig. 2 Data aggregation in WSN

3.3.1. Stage 1

The final destination sends a sparse seed vector $U(u_i), \{i = 1, 2, \ldots, N\}$ with a sparse space △ to all CHs. The location of each CH in this seed vector is identified by its position on the main routing tree.

3.3.2. Stage 2

Starting from its assigned position in the seed vector, the i^{th} CH sends N_i values that rely on the neighbouring nodes N_i. Afterwards, the cluster leader obtains a sparse seed vector and creates its associated sub-matrix $M_i \times N_i$.

3.3.3. Stage 3

CH are not Cluster Leader (CH) nodes transmit their information packets to the CHs. The CHs then transform the obtained packet into M_i using the equation $y_i = \varphi_i x_i$.

3.3.4. Stage 4

The CHs pass the M to the final destination through the identified path.

3.3.5. Stage 5

The final destination assembles the complete measurement matrix using all the values from the seed vector $U(u_i), \{i = 1, 2, \ldots, N\}$ and reconstructs the actual data based on the acquired data $y = [y_1, y_2, \ldots, y_{nc}]$ using the sparse sensing reconstruction algorithm.

3.4. Energy Consumption

As mentioned earlier, non-cluster leader (non-CH) nodes send their data measurements to their designated cluster leaders. The energy used by a cluster node is known as the $P_{intra-cluster}$. In the subsequent phase, CHs collect their specific measurement values ($y_i = \varphi_i x_i$) from the data received by their neighbors and then pass these values on to the final destination. The energy utilized by a cluster node is referred to as P_{BS}, and the total energy expenditure can be represented as follows:

$$P_{total} = P_{intra-cluster} + P_{BS} \tag{8}$$

$$P_{intra-cluster} = N_c \left(\frac{N}{N_c} - 1\right) E[r^{\propto}] \tag{9}$$

This study focuses on a situation where r represents the distance from a shared point to the center of its cluster, and α stands for the path loss exponent. For this research, this work has chosen $\alpha = 2, which$ allows us to determine the expected value $E[r^2]$.

$$E[r^2] = \int\int x^2 + y^2)\, \rho(x, y) dxdy \tag{10}$$
$$= \int\int r'^2\, \rho(r', \theta) r' dr' d\theta$$

The arrangement of nodes is described by $\rho(r, \theta)$ is also considered. Additionally, assume that each cluster is shaped like a circle with a radius R, where R is calculated as R/N_c, where N_c represents the cluster count, and N represents the node count. The nodes within these clusters are assumed to be spread out evenly.

$$E[r^2] = \frac{1}{(\pi L^2/N_c)} \int_{\theta=0}^{2\pi} \int_{r'=0}^{R} r'^3\ dr'\ d\theta = \frac{L^2}{2N_c} \tag{11}$$

$$P_{intra-cluster} = \left(\frac{N}{N_c} - 1\right) \frac{L^2}{2} \tag{12}$$

The energy used in transmitting data between clusters is defined as follows:

$$P_{BS} = \sum_{i=1}^{N\_c} NoH\ (i) * R^2 * M(i) \tag{13}$$

In this scenario, $M(i)$ shows that the count of measurement values linked to the i^{th} cluster, while R^2 stands for the energy cost for each transmission. For analysis, this work operates under the assumption that all clusters have the same size. Previous research suggests that the number of M needed for each cluster is directly related to its size. Therefore, the work can simplify Equation (14) as follows:

$$P_{BS} = R^2 * \frac{M}{N_c}\sum_{i=1}^{N_c} NoH\ (i) \tag{14}$$

In this equation, M denotes the measurement values count collected across the system and N_c represents the cluster count. It can be reformulated as Equation (15):

$$P_{BS} = NoH_{avg} * R^2 * M \tag{15}$$

Where, NoH_{avg} indicates the average number of transmission steps taken.

3.5. Data Transmission Model

The distance at which CH can communicate shows a direct relationship with the network's energy use. Specifically, the number of transmission steps or hops is directly associated with the distance at which transmission can occur. As the distance increases, the cluster's head can interact with more other CHs, reducing the total number of hops needed. In the study, this work creates a network with 2000 nodes using clustering methods like K-means. This work systematically changes the distance at which communication can occur from $R = \{10, 12, 14, 16, 18, 20\}$ to see how changes in distance affect the sum of clusters. Figure 3 shows that the total number of hops changes with changes in distance. Typically, the distance is measured in meters (m) or seconds (s), and this work will not include unit specifications. Figure 3 displays the association between the total number of hops and changes in distance, while Figure 4 compares the total energy used as distance changes. Figure 4 uses K-means before data is combined and uses clustering before data is combined. $RTMA^2$ incorporates a trust evaluation system that assigns trust values to nodes based on their behavior and historical interactions. This trust mechanism ensures that only trusted nodes participate in data aggregation, reducing the likelihood of compromised nodes being selected as aggregators. To mitigate risks from a single compromised aggregator, $RTMA^2$ employs redundant aggregation by using multiple aggregators to handle the same data. This way, even if one aggregator is compromised, the final aggregated result can still be verified and validated through comparison with results from other aggregators.

4. Numerical Results

The section explores the numerical results obtained from simulations and evaluations of the suggested data aggregation technique. The simulation is done in MATLAB 2020a environment and executed in Intel i7 process with 2GB RAM. This work thoroughly evaluates the effectiveness of various strategies by measuring the total data sent by various nodes in that network region using a measure called $\triangle = 2$. The analysis includes six different strategies: (a) the $K-$ means clustering approach with an RSS matrix, (b) the clustering technique also using an RSS matrix, (c) the K-means clustering approach with a Gaussian matrix, (d) the clustering technique with a Gaussian matrix, (e) the K-means clustering strategy that doesn't use compressive sensing, and (f) the clustering strategy that also doesn't use compressive sensing.

This work uses the node count that varies from five-hundred to thousand five-hundred, sets the transmission range at ten, and defines the compression ratio as$\rho = M/N$. Figure 5 illustrates the evaluation of data packet transmission for different strategies at $\rho = 0.2$. It does the same for $\rho = 0.1$. The compression ratios of $\rho = 0.1\ and\ \rho = 0.2$ are set as reference points; the ideal compression ratio for practical use will depend on the specific needs, with higher ratios like $\rho = 0.2$ being more suitable for important applications. Additionally, the work displays how the network's life cycle changes as the number of nodes increases, specifically at a compression ratio of $\rho = 0.1$. The provided values present the network's life cycle progression with more nodes at a compression ratio of $\rho = 0.2$. The visual data clearly show that using compressive sensing significantly extends the network's lifespan compared to using a Gaussian random matrix, as it minimizes the data packets collected, thus lowering the frequency of network operations.

In this research, this work set up a network with 2000 nodes and a specified size of $L = 100$. The initial step involves dividing the network into segments using either K-means clustering, which leads to the formation of clusters called N_c. Following this, a data aggregation technique is applied to evaluate the total energy consumption across the network. Within the defined sensing area, this work designates a central node to act as the sink node. With a total of 500 measurements M available, the goal is to achieve an error threshold of 0.1. To achieve this, this work adjusts the size of the cluster top by changing the transmission range of the nodes. Examine a range of transmission distances $R = [50, 30, 25, 22, 18, 14, 11]$, which correspond to different CH sizes $N_c = [10, 50, 100, 200, 300, 400, 500]$. The first step involves assessing the energy usage linked to each setup. This work carries out tests within clusters, randomly choosing between a sparse matrix and a Gaussian matrix for a side-by-side comparison with different random sparse matrices. These figures display the combined energy use within clusters, including the CH; (b) present the energy consumption patterns using k-means clustering and random sparse matrices; (c) represent the use of K presents with Gaussian that are random matrices; (d) highlight the use of k-means and Gaussian are random matrices.

More clusters lead to lower energy use within clusters. The energy spent on data transfer between clusters is a significant factor. It also shows that sparse matrices use less energy than Gaussian matrices, mainly due to their higher number of zeros. As the number of clusters increases, this work explores various situations: (a) applying K-means with a larger CH count, (b) using k-means clustering with clusters having more leaders, (c) applying K-means with Gaussian matrices and (d) using Gaussian matrices.

Furthermore, as more clusters are added, detail situations like (a) K-means with clusters having a greater number of leaders, (b) k-means with clusters having more leaders, (c) K-means with Gaussian matrices, and (d) k-means with Gaussian matrices. Figure 4 shows a decrease in overall energy use for data communication as the cluster count increases.

This work presents situations for (a) uniform clustering and (b) K-means. Figure 4 shows that this is the pattern of overall energy consumption throughout the network. The study indicates that using multi-hop routing for clusters significantly lowers the network's energy usage, especially with a large number of clusters. a) This situation showcases multi-hop routing used with K-means in clusters. b) This situation illustrates multi-hop routing with k-means in clusters. c) This situation focuses on the use of K-means. d) This situation is about using clustering.

Additionally, this work conducts tests to evaluate the total energy consumption of WSNs using different approaches. In the examination of total energy use, a unique scenario is shown in Figure 8, which arises from the network's dynamic setup with network convergence.

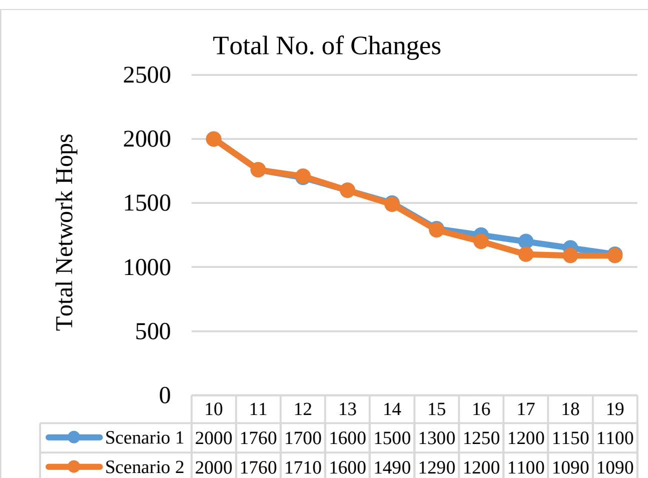

Fig. 3 Total no. of changes vs. Total network hops

This flexibility includes circumstances in which nodes move between clusters or leave the network, demonstrating the WSN's ability to independently adapt to its application requirements. This adaptability has an impact on data collection as well, particularly as additional nodes are added. This paper makes the assumption that certain nodes are added at the beginning of the system structure, with 50 nodes added every two cycles, in order to investigate an extreme case. Comparing the suggested technique to alternative methods both within and between clusters, the results demonstrate a considerable reduction in the network's overall energy consumption.

Furthermore, Figure 6 compares the effects of several data-gathering techniques that illustrate the strategy on the network's duration. When compared to alternative approaches, the data indicates that the methodology prolongs the network's operating life. The algorithm incorporates a method for reducing inter-cluster energy consumption and is created with a focus on the internal dynamics of the clusters utilizing reliable methods. Finding an energy balance when choosing CH is a crucial problem in WSN cluster formation, which the k-means technique effectively solves.

Clustering becomes more important as the number of nodes in the network grows, but it is also necessary to compare the method to other clustering strategies while taking into account the related costs, which are based on bandwidth usage and energy expended on additional communication. RTMA[2] differs from other techniques by facilitating more security, trust, load balancing, and randomness in association to improve the clustering procedure. Its energy-aware mechanisms, combined with trust-based selection and reduced overhead, lead to significant improvements in energy efficiency, making it better suited for long-term, energy-constrained applications in wireless sensor networks.

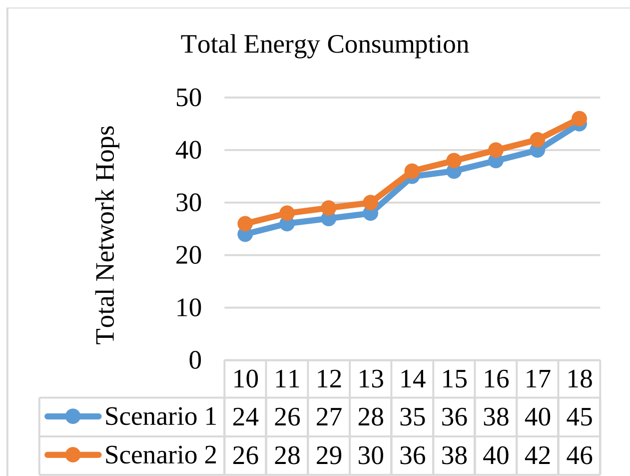

Fig. 4 Total energy Consumption vs. Total network hops

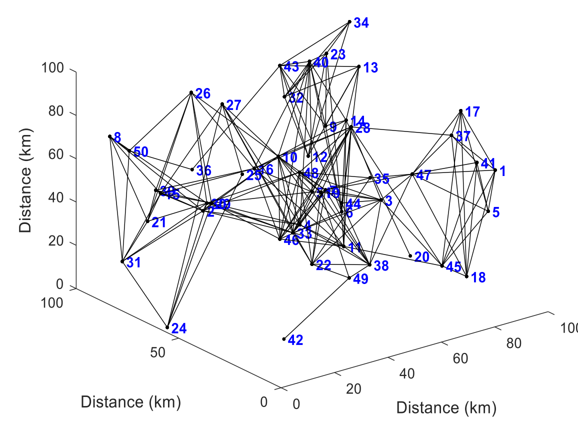

Fig. 5 Node distance measure

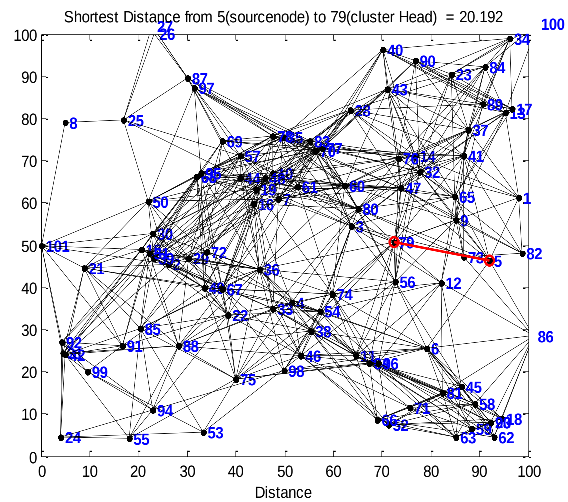

Fig. 6 Shortest distance from source to CH

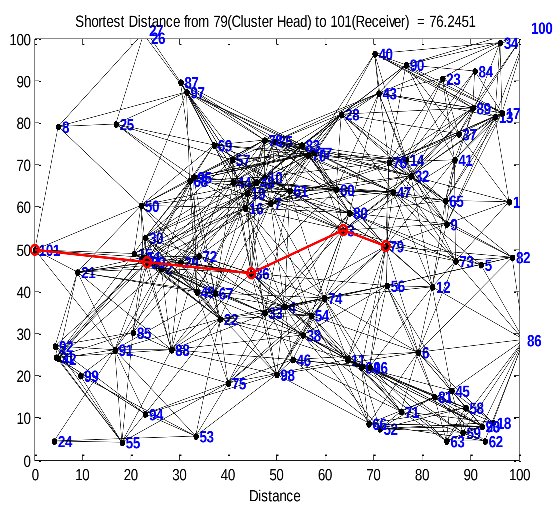

Fig. 7 Shortest distance measure for cluster aggregation

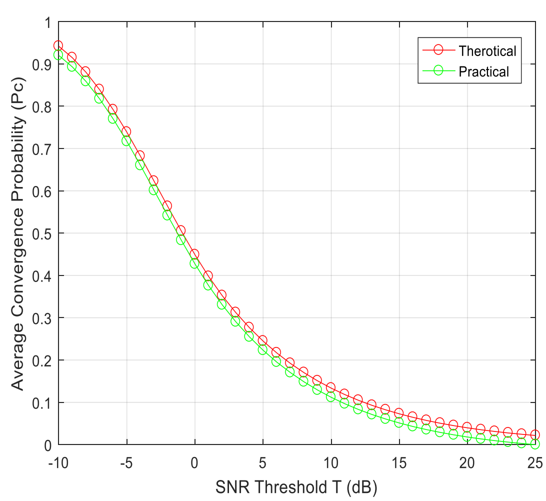

Fig. 8 Average convergence probability of CH

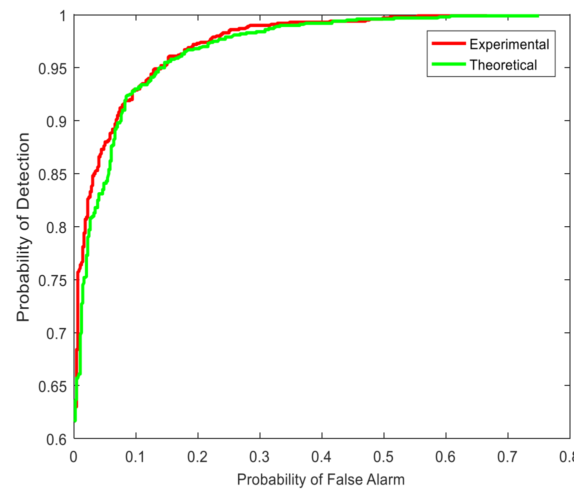

Fig. 9 Probability of trust detection rate

Table 1 depicts the assessment of the suggested RTMA^2 with the prevailing approaches like AOA, WOA, and PSO. The energy consumed by the proposed model is 26J, 20.9 ms, 96.6% throughput, 264.9 rounds of NLT, 94.8% PDR, and 98.9% reliability, which is substantially more than AOA, particle swarm optimization, and WOA. If RTMA^2's energy efficiency scores a mean of 0.8 while the other algorithm scores a mean of 0.7 with a p-value of 0.01, the results are statistically significant, suggesting that RTMA^2 outperforms the alternative method. A small p-value that is less than 0.05 represents that RTMA^2's improvement over the competing method is statistically significant.

Table 1. Outcomes analysis

Metrics	RTMA^2	AOA	WOA	PSO
Reliability (%)	98.9	96.15	92.9	90.9
NLT (rounds)	264.9	222.46	199.9	177.5
Throughput (%)	96.6	76.13	73.5	67.8
ECP (J)	26	6231.01	6489.13	7164.1
PDR (%)	94.8	90.35	88.4	84.5
Delay (ms)	20.9	27.13	30.6	38.1

5. Conclusion

This work introduces a novel Random Trusted Multi-Aggregation Algorithm (RTMA^2) strategy designed to enhance data gathering in WSN, significantly reducing energy consumption within the network. This process starts with the sink node sending out a limited initial signal to the CHs. Each CH uses this sparse signal to create the necessary measurement matrix and derive the necessary measurements that are used for values using random that are sparse in compressive sensing. Afterwards, these size values are passed back to other sink nodes through a multi-hop routing tree linking the clusters. The sink node then applies a specific compressive sensing algorithm to renovate the actual signal.

The analysis centres on the power efficiency of this approach within the network, how CH size impacts the energy needed for inter-cluster communications, and the effect of CH size on the overall energy utilization of the network. The findings depict that this technique successfully leads to a reduction in the amount of energy used by the network. In the future, Clustering algorithms can be enhanced to dynamically adjust cluster head selections based on nodes' energy levels, proximity, and data quality. Adaptive clustering can reduce the burden on specific nodes, extending the network's lifespan. The effectiveness of data aggregation can be significantly enhanced by employing adaptive algorithms. Traditional data aggregation techniques often use fixed parameters, but environmental factors and network dynamics, such as node mobility, energy depletion, or varying data generation rates, can affect their performance. Adaptive algorithms adjust in real time based on network conditions, improving overall efficiency.

References

[1] Raniyah Wazirali et al., "Secure Watermarking Schemes and Their Approaches in the IoT Technology: An Overview," *Electronics*, vol. 10, no. 14, pp. 1-28, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Khalid Haseeb et al., "RTS: A Robust and Trusted Scheme for IoT-Based Mobile Wireless Mesh Networks," *IEEE Access*, vol. 8, pp. 68379-68390, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Khattak Hasan Ali et al., "Cross-Layer Design and Optimization Techniques in Wireless Multimedia Sensor Networks for Smart Cities," *Computer Science and Information Systems*, vol. 16, no. 1, pp. 1-17, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Avishek Banerjee et al., "Design of Energy Efficient WSN Using a Noble SMOWA Algorithm," *Computers*, *Materials & Continua*, vol. 72, no. 2, pp. 3585-3600, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Zeeshan Ahmad et al., "Anomaly Detection Using Deep Neural Network for IoT Architecture," *Applied Sciences*, vol. 11, no. 15, pp. 1-19, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Laxminarayan Sahoo et al., "Optimization of Data Distributed Network System under Uncertainty," *Discrete Dynamics in Nature and Society*, vol. 2022, no. 1, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Chirihane Gherbi, Zibouda Aliouat, and Mohamed Benmohammed, "An Adaptive Clustering Approach to Dynamic Load Balancing and Energy Efficiency in Wireless Sensor Networks," *Energy*, vol. 114, pp. 647-662, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Palash Dutta, "Multi-Criteria Decision Making under Uncertainty via the Operations of Generalized Intuitionistic Fuzzy Numbers," *Granular Computing*, vol. 6, pp. 321-337, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] A. Cagri Tolga, I. Burak Parlak, and Oscar Castillo, "Finite-Interval-Valued Type-2 Gaussian Fuzzy Numbers Applied to Fuzzy TODIM in a Healthcare Problem," *Engineering Applications of Artificial Intelligence*, vol. 87, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Elaheh Moharamkhani et al., "Multiobjective Fuzzy Knowledge-Based Bacterial Foraging Optimization for Congestion Control in Clustered Wireless Sensor Networks," *International Journal of Communication Systems*, vol. 34, no. 16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sonam Lata et al., "Fuzzy Clustering Algorithm for Enhancing Reliability and Network Lifetime of Wireless Sensor Networks," *IEEE Access*, vol. 8, pp. 66013-66024, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Damien Wohwe Sambo et al., "Optimized Clustering Algorithms for Large Wireless Sensor Networks: A Review," *Sensors*, vol. 19, no. 2, pp. 1-27, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Y.U. Xiu-wu et al., "A Clustering Routing Algorithm Based on Wolf Pack Algorithm for Heterogeneous Wireless Sensor Networks," *Computer Networks*, vol. 167, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] S.R. Deepa, and D. Rekha, "Cluster Optimization in Wireless Sensor Networks Using Particle Swarm Optimization," *Social Transformation – Digital Way: 52nd Annual Convention of the Computer Society of India*, Kolkata, India, pp. 240-253, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] Chun-Ho Chen, "A Novel Multi-Criteria Decision-Making Model for Building Material Supplier Selection Based on Entropy-AHP Weighted TOPSIS," *Entropy*, vol. 22, no. 2, pp. 1-23, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Laxminarayan Sahoo, "Some Score Functions on Fermatean Fuzzy Sets and its Application to Bride Selection Based on TOPSIS Method," *International Journal of Fuzzy System Applications*, vol. 10, no. 3, pp. 18-29, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Dinesh Kumar Kotary, and Satyasai Jagannath Nanda, "Distributed Robust Data Clustering in Wireless Sensor Networks Using Diffusion Moth Flame Optimization," *Engineering Applications of Artificial Intelligence*, vol. 87, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Syed Fawad Hussain, and Shahid Iqbal, "CCGA: Co-Similarity Based Co-Clustering Using Genetic Algorithm," *Applied Soft Computing*, vol. 72, pp. 30-42, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] K.B. Gurumoorthy et al., "A Novel Clustering Method for Fault Recovery and Routing in Mobile Ad-Hoc Networks," *International Journal of Communication Systems*, vol. 34, no. 15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Shunchao Zhang, "Riemannian Distance-Based Fast K-Medoids Clustering Algorithm for Cooperative Spectrum Sensing," *IEEE Systems Journal*, vol. 16, no. 1, pp. 880-890, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Saima Zafar, A. Bashir, and S.A. Chaudhry, "Mobility-Aware Hierarchical Clustering in Mobile Wireless Sensor Networks," *IEEE Access*, vol. 7, pp. 20394-20403, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[22] Supriyan Sen et al., "Wireless Sensor Network Lifetime Extension via K-Medoids and MCDM Techniques in Uncertain Environment," *Applied Sciences*, vol. 13, no. 5, pp. 1-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Proshikshya Mukherjee et al., "Recommended System for Cluster Head Selection in a Remote Sensor Cloud Environment Using the Fuzzy-Based Multi-Criteria Decision-Making Technique," *Sustainability*, vol. 13, no. 19, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Bryan Raj et al., "A Survey on Cluster Head Selection and Cluster Formation Methods in Wireless Sensor Networks," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-53, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Rakesh Kumar Yadav, and R.P. Mahapatra, "Energy Aware Optimized Clustering for Hierarchical Routing in Wireless Sensor Network," *Computer Science Review*, vol. 41, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] Pramod Singh Rathore et al., "Energy-Efficient Cluster Head Selection through Relay Approach for WSN," *The Journal of Supercomputing*, vol. 77, pp. 7649-7675, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] Amir Abbas Baradaran, and Keivan Navi, "HQCA-WSN: High-Quality Clustering Algorithm and Optimal Cluster Head Selection Using Fuzzy Logic in Wireless Sensor Networks," *Fuzzy Sets and Systems*, vol. 389, pp. 114-144, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Abel García-Nájera, Saúl Zapotecas-Martínez, and Karen Miranda, "Analysis of the Multi-Objective Cluster Head Selection Problem in WSNs," *Applied Soft Computing*, vol. 112, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] Mohammad Al-Quraan et al., "Edge-Native Intelligence for 6G Communications Driven by Federated Learning: A Survey of Trends and Challenges," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 7, no. 3, pp. 957-979, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Ihsan Ullah, Hee Yong Youn, and Youn-Hee Han, "An Efficient Data Aggregation and Outlier Detection Scheme Based on Radial Basis Function Neural Network for WSN," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[31] G. Saranraj, K. Selvamani, and P. Malathi, "A Novel Data Aggregation Using Multi Objective Based Male Lion Optimization Algorithm (DA-MOMLOA) in Wireless Sensor Network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 5645-5653, 2022. [CrossRef] [Google Scholar] [Publisher Link]