

# Hybrid Swarm-Autoencoder Model for Energy-Efficient Robust Clustering in Wireless Sensor Networks

Syed Noor Syed , Geethanjali Nellore 

Department of Computer Science and Technology, Sri Krishnadevaraya University, Anantapur, India

Corresponding author: Syed Noor Syed (syednoor.s@gmail.com)

## Editorial Record

**First submission received:**  
August 29, 2025

**Revisions received:**  
November 14, 2025  
February 8, 2026  
March 16, 2026

**Accepted for publication:**  
March 21, 2026

**Academic Editor:**  
Zdenek Smutny  
Prague University of Economics  
and Business, Czech Republic

This article was accepted for publication  
by the Academic Editor upon evaluation of  
the reviewers' comments.

**How to cite this article:**  
Syed, S. N., & Nellore, G. (2026). Hybrid  
Swarm-Autoencoder Model for Energy-  
Efficient Robust Clustering in Wireless  
Sensor Networks. *Acta Informatica  
Pragensia*, 15(2), 416–439.  
<https://doi.org/10.18267/j.aip.314>

**Copyright:**  
© 2026 by the author(s). Licensee Prague  
University of Economics and Business,  
Czech Republic. This article is an open  
access article distributed under the terms  
and conditions of the [Creative Commons  
Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).



## Abstract

**Background:** Environmental monitoring and data collection critically depend on wireless sensor networks (WSN). However, the limited battery life of these sensor nodes significantly affects the operational lifetime of the network. This often results in challenges related to energy efficiency and clustering tasks due to dynamic network topologies and limited resources.

**Objective:** This paper aims to improve clustering efficiency and extend the network lifetime of WSN by proposing a hybrid model.

**Methods:** The proposed framework was evaluated in terms of energy consumption, network lifetime, Packet Delivery Ratio (PDR), and clustering efficiency through extensive simulations. Modified Salp Swarm Algorithm (MSSA)-Deep Stacked Sparse Autoencoder (DSSA)-K-means is a promising solution for next-generation WSN in smart environments and industrial Internet of Things (IoT) applications because experimental results show that it achieves higher clustering accuracy, energy economy, and robustness against network failures than current methods.

**Results:** Extensive simulations were conducted to evaluate the proposed framework based on energy consumption, network lifetime, PDR, and clustering efficiency. The experimental results demonstrate that the proposed MSSA-DSSA-K-means model achieves higher clustering accuracy, energy economy, and robustness against network failures compared to existing methods.

**Conclusion:** The paper concludes that this hybrid model is a promising solution for next-generation WSN in smart environments and industrial IoT applications.

## Index Terms

Wireless sensor networks; Energy-efficient clustering; Modified salp swarm algorithm; Deep stacked sparse autoencoder; K-means clustering; Cluster head selection.

## 1 INTRODUCTION

WSN have become the cornerstone of emerging data-driven environments such as smart cities, healthcare systems, industrial automation, and the IoT backbone. A WSN is a network of spatially distributed autonomous nodes with sensing, computational, and communication capabilities. Nevertheless, inherent WSN constraints (energy-limited, dynamic topologies, resource heterogeneity, etc.) require mechanisms to perform efficient data delivery over the network and maintain a longer network lifetime (Nathiya et al., 2025; Anguraj et al., 2024). To overcome these issues, energy-efficient clustering has been proposed, where nodes are partitioned into clusters and a particular node is selected as the Cluster Head (CH), which is responsible for the aggregated fusion and transmission of data towards the sink.

Classical algorithms, such as Low Energy Adaptive Clustering Hierarchy (LEACH) and Hybrid Energy Efficient Distributed Clustering (HEED), which are the basis of these and other similar algorithms, do not perform well on heterogeneous or large-scale deployments, mainly because of the static nature of their parameter tuning and their inability to adapt to actual operating conditions (Bi et al., 2019; UmaRani et al., 2025).

A review of current research trends is presented, which indicates hybridization of optimization algorithms and Machine learning (ML) as promising paradigms for improving cluster member selection, CH selection, fault detection and secure routing. The presented a hybrid ML-based clustering framework with fault tolerance, which maximizes the energy and data management efficiency in WSN-IoT applications". In another work, the idea of clustering and fuzzy logic techniques combined with ML was used to prolong the network's lifetime by updating the cluster parameters based on the environmental conditions (Anguraj et al., 2024). Swarm intelligence and nature-inspired algorithms and optimization methods, including Ant Colony Optimization (ACO), Grey Wolf Optimization (GWO), Particle Swarm Optimization (PSO), and Glow Worm Swarm Optimization (GSO), have been proven to be humble contributions to solving multi-objective cluster problems (Amutha et al., 2022; Reddy et al., 2021; Malisetti & Pamula, 2022). These algorithms dynamically balance exploration and exploitation, providing energy efficient paths and reliable CH selection (Reddy et al., 2021; Almusawi et al., 2024). An optimization-based hybrid algorithm with static and mobile sink strategies mitigates energy holes and enhances the data throughput (Amutha et al., 2022).

Swarm intelligence and optimization-based clustering approaches have showed great potential to improve the energy efficiency and load balancing of WSNs. For example, Wang et al. (2020) designed a clustering routing protocol based on the chaotic genetic algorithm to improve the selection of cluster-heads and balance the energy consumption of nodes, which can prolong the network lifetime and improve the routing performance. The results further show the efficiency of metaheuristic optimization techniques in solving energy-constrained routing and clustering problems of wireless sensor networks.

In addition, deep learning architectures such as DSSA and Convolution Neural Networks (CNNs) have been combined with unsupervised clustering (e.g., K-Means) to automatically learn feature representations and to achieve adaptive network performance in nonstationary contexts (Bi et al., 2019; Madkar et al., 2023). A hybrid ML model along with an enhanced Salp Swarm Optimization (SSO) enhances packet delivery and minimizes the latency in dense WSN (UmaRani et al., 2025). Security and trust management are top-level issues in WSN applications, particularly adversarial environments in mission-critical applications. Bio-inspired clustering supports tamper-proof communication and energy-efficient transmission via blockchain-tailored routing for edge-assisted WSN (Prasad & Periyasamy et al., 2023). Hybrid cryptographic systems enhance lightweight and secure communication protocols (Selvi et al., 2024; Jalalinejad et al., 2024). One other dimension to be considered is energy harvesting and renewables integration, in which more recent models have considered optimisation of the routing in the presence of fluctuating energy, making it possible to have a sustainable operation's (Jayachandran & Devi, 2024; Vanitha et al., 2023; Jing et al., 2024) Advanced hybrid protocols such as Energy Efficient Routing Cluster Generation using Hybrid Harrys Hawks Optimization Algorithm (EER-CGHHOA) and Salp Swarm PSO show how multi-hop clustering together with dynamic optimization can cope with real-time constraints and achieve maximum throughput. Given this progress, the integration of machine learning, hybrid optimization, and bio-inspired computing increasingly restructures WSN models.

The work in (Jamaesha et al., 2024; Sreekanth et al., 2023). embarks on the further development of intelligent clustering schemes, which not only desire to improve energy efficiency but also wish to include security, fault tolerance, adaptive learning and scalability, which are considered the pillars for future wireless sensor systems. The contribution of this work as follows:

- To develop a hybrid optimization framework for the design of an MSSA-DSSA-K-means framework to improve energy-efficient clustering in WSN.
- Improvement of clustering efficiency and accuracy Deep feature extraction and dimensionality reduction using DSSA will help improve clusters based on K-means clustering, minimize intra-cluster communication overhead and maximize cluster development.
- Balancing the network energy usage will help integrate MSSA for the best CH choice.
- To improve scalability and network lifetime adaptive clustering helps lower energy usage and extend the lifetime of networks.

- Existing Algorithms: Compared with benchmarks against advanced clustering and optimization models such as Entropy-based Analytic Hierarchy Process Weighting Technique for Order of Preference by Similarity to Ideal Solution(EAHPW-TOPSIS), Drawer Algorithm-Based Energy Efficient Cluster Head Selection (DA-EECHS), Genetic-Based Energy Efficient Clustering (GEEC), and Energy Aware Distance-Based Cluster Head Selection and Routing (EADCR) proposed the MSSA-DSSA-K-Means.
- We analysed the performance indices using MATLAB-based simulations and evaluated the suggested model with respect to energy economy, PDR, throughput, and latency.

**Problem statement:** Limited battery resources, dynamic topologies, and imbalanced energy consumption among sensor nodes make WSN difficult to use for energy-efficient data transmissions. Traditional clustering methods often have issues such as excessive communication within the cluster, choosing the wrong CH, and not using energy efficiently, which can make the network less stable and shorten its lifespan. In this study, a MSSA was coupled with a DSSA and K-means clustering to solve these constraints. DSSA is used for dimensionality reduction and feature extraction to improve cluster accuracy. MSSA ensures optimal CH selection, achieves energy balance, and reduces latency, while K-Means maximizes cluster development, reducing communication overhead. Owing to non-uniform energy distribution, redundant data transmission, and ineffective routing protocols, robust and scalable clustering remains a major challenge for WSN, despite improvements in clustering methods. This effort seeks to raise the general energy efficiency, lengthen the network lifetime, and strengthen the clustering robustness.

Benchmarking the proposed MSSA-DSSA-K-Means framework against state-of-the-art clustering and optimization models, including Cluster based Ant Colony Optimization NETWORK (CACONET), Improved Cluster based Moth Flame Optimization (ICMFO), Firefly Algorithm – Optimized Link State Routing(FA-OLSR), and Adaptive Neuro Fuzzy Clustering – Quasi Chance Seed Optimization Routing(ANFC-QGSOR), shows its superiority in energy-aware clustering for next-generation WSN in smart environments and industrial IoT applications.

## 2 LITERATURE SURVEY

Energy-efficient routing and clustering algorithms in WSN have attracted the attention of researchers, especially where the demand for the IoT is increasing in the domains of smart cities and industry. The related literature reviews advanced optimization techniques and hybrid clustering models for energy constraints, network lifetime performance, and communication effectiveness in WSN.

Sharma et al. (2024) continued to explore the issue of long-term energy limitation in Industrial Wireless Sensor Networks (IWSN). Control packet volume in IWSN may continue to cause a reduction in network longevity. The authors developed the Novel Hybrid Ant Routing using Swarm Optimization (NHARSO-IWSN) protocol, which is an adaptive fuzzy-inference and reptile search optimization model designed to minimize the amount of energy consumed unnecessarily by the sensor nodes. The results from their simulations demonstrate that the NHARSO-IWSN protocol significantly reduces energy use while also increasing anomaly detection accuracy, which shows that hybrid fuzzy-optical models are very suitable for environments that have limited power available on sensor nodes. The research demonstrates adaptive rule-based systems may be used as complementary techniques to optimization algorithms in industrial monitoring.

Liu et al. (2024) have been able to improve the lifespan of clusters in a WSN, especially during their start-up phase when many sensor nodes die prematurely. To improve the lifespan of clusters, they proposed a dual CH structure that has been improved by a hybrid Grey Wolf and Whale Optimization Algorithm (GWOA-CH). The idea behind this is to frequently elect new CHs so that one node does not consume all the high-energy communication resources. A comparison was made between GWOA-CH and some other algorithms, such as a modified genetic algorithm, a non-dominated crowding-based opposition-based genetic algorithm, and a modified multi-objective random search optimizer for routing, and it appeared that GWOA-CH had a positive effect on the amount of energy saved and on the workload distribution, which both contribute to an extended life span of the network. The use of hybrid swarm intelligence may be a viable solution to the chronic load imbalance problem in traditional clustering.

Wang et al. (2024) address a stab at the "hotspot" problem of CH's premature death because of excessive traffic. The authors developed an Unequally Clustered Routing Protocol with Multi-Hop Threshold Distance (UCRTD), which determines relay path based upon residual energy and distance. UCRTD uses multi-hop routes to do intra-cluster communications and does intelligent cluster head selection for inter-cluster communications, so it has better load balancing. The experiments have shown that UCRTD supports large numbers of nodes in a WSN deployment and

significantly increases the longevity of WSN networks, thereby reinforcing the benefit of using dynamically determined threshold routing.

Yassine et al. (2019) also investigated the revolution brought to society by the IoT, in which WSN function as the groundwork of context-aware intelligent systems. When scaling up to domain applications such as smart agriculture and environmental monitoring in the IoT, the performance and energy sustainability of WSN is a challenge. To address these issues, the authors designed a fuzzy logic-based clustering algorithm to minimize the energy consumption and enhance the scalability of the network and data-link layers. It is shown through simulations that the performance of this algorithm is more favourable in cases where a considerable number (up to approximately half) of nodes out of the network have empty energy. The method is robust and flexible and is a potential candidate for large-scale IoT deployment.

Zachariah et al. (2022) proposed two new energy-efficient clustering algorithms, Hybrid Optimization Clustering using Krill Herd and Cuckoo Search (HOCK) and Hybrid Energy-efficient Clustering using Krill herd and Cuckoo search (HECK), to ensure homogeneous and heterogeneous scenarios, respectively, to extend the lifetime of WSN for both homogeneous and heterogeneous environments. Swarm intelligence algorithms, such as Krill Herd Optimization (KHO) and Cuckoo Search (CS) are used to find the best spots for cluster heads and centroids to reduce the distance within clusters and save energy. HOCK is tested under different base station locations and node densities, and HECK considers multi-tier heterogeneity. Simulation results show that both schemes achieve a notable improvement in network lifetime, and thereby HOCK has a better performance than classic clustering algorithms.

Godi et al. (2025) presented a hybrid optimization method to enhance the energy efficiency and routing performance of a WSN. The power of Zone-based Optimization Algorithm (ZOA) and Bacterial Foraging Optimization (BFO) is combined to present the most recent model of the Hybridized Zebra and Bitterling fish optimization algorithm (HZ-BFOA) that synergizes the exploration strategy of ZOA with the adaptive searching capabilities of feral fish, or BFO. This mixed scheme is utilized to determine the best Cluster Head (CH) nodes and routing paths. Multi-objective optimization reduces energy consumption and data latency and simultaneously improves throughput at the same time. A Quality of Service (QoS) performance analysis of the HZ-BFOA model also indicates an increase in throughput of up to 89%, while energy reduction is considerable and less data packet loss demonstrates the potential for a reliable solution for mission-critical WSN applications in defence and public safety.

Kumar et al. (2022) addressed the problem of detecting malicious nodes and ensuring secure data transmission in resource-constrained WSN within the IoT. In this paper, we propose an Improved Deep Convolutional Neural Network (IDCNN) model for the identification and isolation of Malicious Nodes (MNs) according to their behaviour and communication. After identifying the Transport Node (TN)s, an Efficient Key Management (EKM) algorithm is used to group the TNs into clusters, whereas the Trust-based Distributed Secure Beaconing Optimization (T-DSBO) algorithm selects the optimum CHs based on residual energy. This hybrid model provides security while also improving energy efficiency by avoiding malicious interference and achieving reliable, energy-aware data transfer, which improves overall network reliability.

Karunkuzhali et al. (2025) present a new clustering and optimization model, named MSGAN-RPOA-DAA-EERP that combines the Multi-Stream Generative Adversarial Network (MSGAN) and Red Panda Optimization Algorithm (RPOA) for effective data fusion and energy-aware communication in WSN. This method refines the parameters of the residual node energy and communication energy cost on nodes and cluster heads, respectively, ultimately leading to significant energy efficiency. The simulation results reveal that the energy consumption is reduced by 20.28%, 27.91%, and 17.53%, respectively, compared to EECOP-PIAS-WSN, IBO-DFNL-EEDA and EC-WSN-EEELCP. The proposed model is found to be robust to real-time energy optimization for WSN and can be employed in long-term deployment in dynamic IoT environments.

Rathee et al. (2021) proposed a QoS-aware, energy-balanced, and secure routing algorithm based on ACO for WSN. One of the proposed algorithms is QoS-aware energy balancing secure routing (QEBSR), which considers the transmission delay, trust metrics, and energy conditions of nodes for secure and efficient path routing. Theoretical analysis and simulation results demonstrate that in terms of throughput, delay, and network lifetime, QEBSR outperforms existing schemes, for example, distributed energy-balanced routing and node-compromise-resistant energy-efficient routing. The experimental results demonstrate that the incorporation of QoS parameters and trust-based heuristics into ACO can significantly enhance the reliability and scalability of WSN routing protocols.

Thekiya and Nikose (2022) suggested using a method that combines Fuzzy C-Means (FCM) and an Artificial Bee Colony (ABC) approach for grouping and selecting CH in WSN. FCM helps determine the number of groups that should be, whereas ABC is used to choose the CH. The FCM method also checks how well the grouping works by examining factors such as energy balance, load distribution, GINI coefficient, distance between clusters, and fair energy use by the sensor nodes. In addition, the ABC algorithm enhances data transmission using an energy-efficient ACO technique. The FCM approach checks the performance of the clustering with the help of load-balancing parameters, such as energy balance, load distribution, GINI coefficient, inter-cluster distance, and unbiased energy consumption by the sensor nodes. In this regard, the ABC algorithm improves data transmission using the energy-aware ACO technique. Furthermore, a dynamic Admission Allotment Scheme (AAS) is proposed to manage intra cluster communications, which greatly reduces packet drops and overhead. The proposed approach brings significant enhancements in the network lifetime, PDR, and throughput compared to non-intelligent traditional approaches.

Bozorgi et al. (2019) suggested a hybrid uneven energy-efficient clustering protocol to reduce the energy consumption in WSN. Their approach employs a hybrid of static and dynamic clustering to minimize the exchange of control messages. An interesting aspect is its neighbor-aware strategy, which arbitrarily selects to communicate with neighbors, thereby contributing to the general protocol overhead reduction. It also presents two essential routing mechanisms such as the assistance-to-cluster-heads mechanism, in which the member nodes provide the CHs with the assistance of data forwarding, and the discretionary licensing mechanism, which prevents incomplete transmissions of the data packet. In addition, it is beneficial to employ one-layer clustering to enhance routing hierarchy and efficiency. The simulation results show that the protocol has better stability, energy balancing, and network lifetime than other protocols.

Sheik Dawood et al. (2025) proposed a new Hybrid Archimedes Arithmetic Optimization Algorithm (HYB-AO-AOA-CHS-WSN) to enhance CHs selection and multipath routing in WSN. By using the merits of the Archimedes optimization algorithm and the arithmetic optimization algorithm, a dual optimization model is proposed to evaluate the sensor nodes according to essential parameters such as transmission delay, remaining energy, node connections, and the distances between clusters and in clusters. The proposed hybrid protocol is simulated in MATLAB and rigorously analysed using key performance metrics, which include network delay, number of alive nodes, throughput, and remaining energy. The results demonstrate that in an energy-constrained WSN environment, the proposed SRD-DCGAN-boosted HYB-AO-AOA technique attained increased throughput, longer node longevity, and lower delay compared to traditional clustering and routing mechanisms. This study highlights the significance of integrated optimization in terms of WSN scalability, reliability, and data delivery efficiency, in addition of problem formulation of literature survey for WSN shown in Table 1.

**Table1.** Problem formulation of literature survey for WSN.

Author (Year)	Methods	Research Gaps	Limitations	Energy Efficiency / Key Outcome
Sharma et al. (2024)	Hybrid Adaptive Network-based Fuzzy Inference System with Reptile Search Optimization (NHARSO-IWSN)	Existing clustering schemes fail to mitigate excessive control packet overhead in industrial WSN	Performance sensitive to node density and mobility variations	Lower energy consumption; improved detection accuracy
Liu et al. (2024)	Dual Cluster Head (Dual-CH) routing with GWOA-CH (Grey Wolf + Whale Optimization)	Early node death and frequent CH re-selection reduce network stability	Scalability in ultra-dense deployments not fully examined	Superior energy-saving capability and balanced CH workloads
Wang et al. (2024)	Unequally Clustered Routing with Multihop Threshold Distance (UCRTD)	Lack of analytical foundation for threshold distance design and load balancing	Threshold computation increases algorithmic complexity	Improved load balancing and extended network lifetime
Yassine et al. (2019)	Fuzzy Logic-based Clustering Algorithm	Scalability challenges in large IoT-driven WSN applications	Does not address real-time communication or mobility	Better performance in scalable and heterogeneous IoT scenarios
Zachariah et al. (2022)	HOCK & HECK using Krill Herd Optimization (KHO) + Cuckoo Search (CS)	Limited adaptive capability in fully heterogeneous environments	Evaluation restricted to specific heterogeneity levels	High energy efficiency and prolonged network lifetime

Author (Year)	Methods	Research Gaps	Limitations	Energy Efficiency / Key Outcome
Godi et al. (2025)	Hybrid Zebra–Bitterling Fish Optimization Algorithm (HZ-BFOA)	Optimal routing and CH selection under high mobility remains difficult	Computational overhead may increase in large networks	Up to 89% throughput; efficient CH selection and routing
Kumar et al. (2022)	IDCNN for malicious node (MN) detection + EKM & t-DSBO for clustering	Security routing and MN isolation previously handled separately	Focused primarily on specific MN attack categories	Effective MN detection; energy-aware secure routing
Karunkuzhali et al. (2025)	MSGAN + Red Panda Optimization (MSGAN-RPOA-DAA-EERP)	Need for improved energy-aware data aggregation	GAN-based model training is computationally expensive	17.53%–27.91% reduction in energy usage compared to baselines
Rathee et al. (2021)	QoS-Aware Secure Routing using Ant Colony Optimization (QEBSR)	Insufficient integration of QoS, security, and energy parameters	No validation under highly dynamic real-time networks	Enhances QoS, security trust metrics, and energy balance
Thekiya and Nikose (2022)	Fuzzy C-Means (FCM) + Artificial Bee Colony (ABC) for CH selection; ACO for routing	Lack of combined clustering–routing optimization in FCM-based models	Assumes static network topology	Enhanced network lifetime, PDR, and throughput
Bozorgi et al. (2019)	Hybrid Static–Dynamic Unequal Clustering with Inter-cluster Routing (HEEC)	CH overloading remains partially unresolved	Overhead reduction may restrict routing flexibility	Balanced energy consumption; improved stability and lifetime
Sheik Dawood et al. (2025)	Hybrid Archimedes + Arithmetic Optimization Algorithm (HYB-AO-AOA)	Lack of unified multi-objective synergy for clustering and routing	No testing under dynamic real-time mobility or failures	Higher throughput, greater residual energy, and reduced delay

### 3 THE ARCHITECTURE OF THE PROPOSED MODEL

The proposed system combines DSSA, K-means clustering, and MSSA to create a resilient and power-aware clustering technique that is suitable for WSN. The architecture consists of three main stages: (1) feature reduction, (2) initial clustering, and (3) cluster head (CH) optimization using a metaheuristic. First, raw sensor data, which are generally high-dimensional and encoded with redundancy, are input into the DSSA to obtain features of interest and remove noise and irrelevant information. This compact encoding significantly decreases the communication overhead resulting from data transmission. The DSSA result is then input into a conventional k-means to conduct pre-treatment clustering of sensor nodes according to feature similarity.

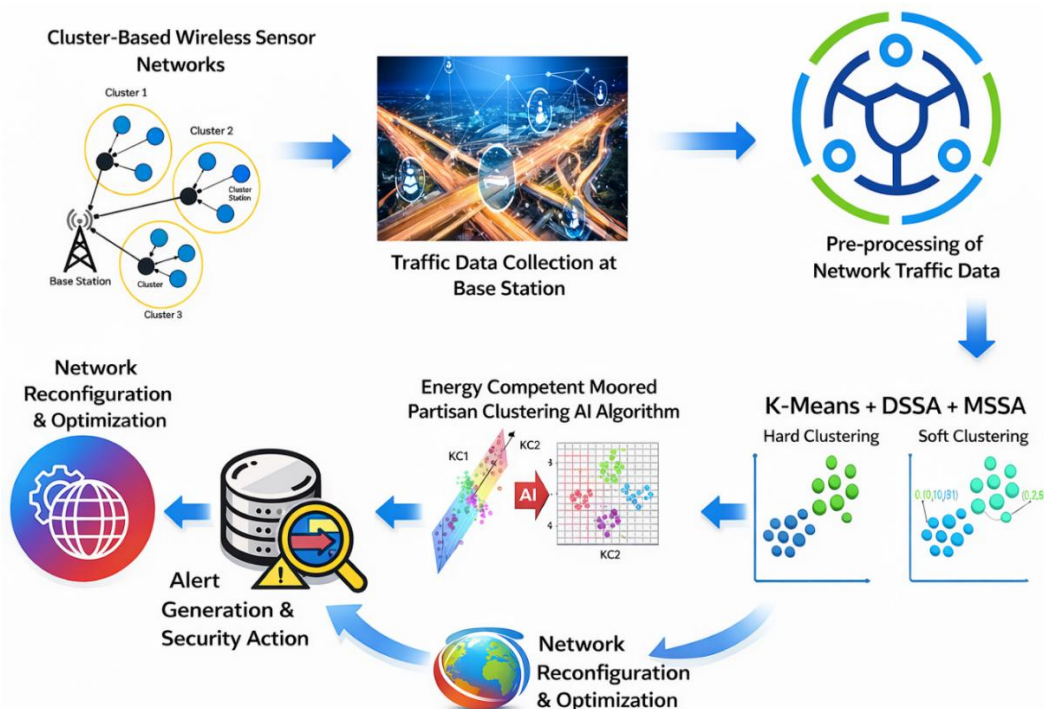


Figure 1. Proposed Block Diagram Cluster based WSN.

Nevertheless, K-means by itself is not sufficient because it is very sensitive to initialization and cannot optimize the energy assignment. To mitigate this, a MSSA is applied for clustering by minimizing a multi-objective cost function that consists of intra-cluster distance, residual energy, and network lifetime. The design guarantees energy-efficient CH rotation and balanced energy consumption by applying adaptation to the rotation record of the CH using the fitness-based optimization technique of the MSSA, resulting in long network survivability and high robustness. figure 1 shows the proposed block diagram flow work operational of cluster WSN. The interaction relationship between components is illustrated, and the data flow among the components is presented in the system flow diagram.

### 3.1 The WSN Model and Assumptions

It includes N static sensor nodes  $S_i = 1, 2, \dots, i$  and  $N_i = 1, 2, \dots, N$  which are randomly deployed in a two-dimensional (2-D) square sensing field  $A \times A$ . The sensor nodes have the following three basic capabilities: (1) sensing environmental or physical parameters (e.g., temperature, humidity, or motion), (2) processing the aggregated data, and (3) wireless communication (e.g., transmitting and receiving data packets). However, these nodes are energy-limited and work under energy-constrained conditions; therefore, they run with constrained battery power, and the operation of nodes is severely affected by throughput. The Base Station (BS) is the centre of information collection and can operate without energy limitations for the entire network procedure. The BS may be placed in the centre of the sensing area, or it is also possible to place the BS outside the deployment area as per the application used. The BS accepts, processes, and stores the data accumulated from sensor nodes, which can also generate command control.

#### 1. Energy Consumption

The WSN energy consumption is modelled using the First-Order Radio Model, which mathematically represents the energy usage during transmission, reception, and aggregation. This model is vital for achieving energy efficiency in networks, as it affects the lifespan of battery-limited sensor nodes.

#### 2. Energy from Transmission ( $E_{tx}(kd)$ )

When a message of size k bits is transmitted by a sensor node at a distance d, the energy consumed is a function of the fact that the transmission is performed at a distance d and  $d_0$ . This threshold separates the two propagation models. For short energy transfers, the power depends on the inverse square law:

$$E_{tx}(kd) = \begin{cases} K \cdot E_{Ele} + k \cdot \epsilon_{fs} \cdot d^2, & \text{if } (d < d_0) \\ K \cdot E_{Ele} + k \cdot \epsilon_{mp} \cdot d^4, & \text{if } (d \geq d_0) \end{cases} \quad (1)$$

Where  $E_{Ele}$  is the transmitter electronics energy consumption per bit, and  $\epsilon_{fs}$  is the free-space path loss amplifier energy, and  $\epsilon_{mp}$  is amplifier energy for the multipath fading model.

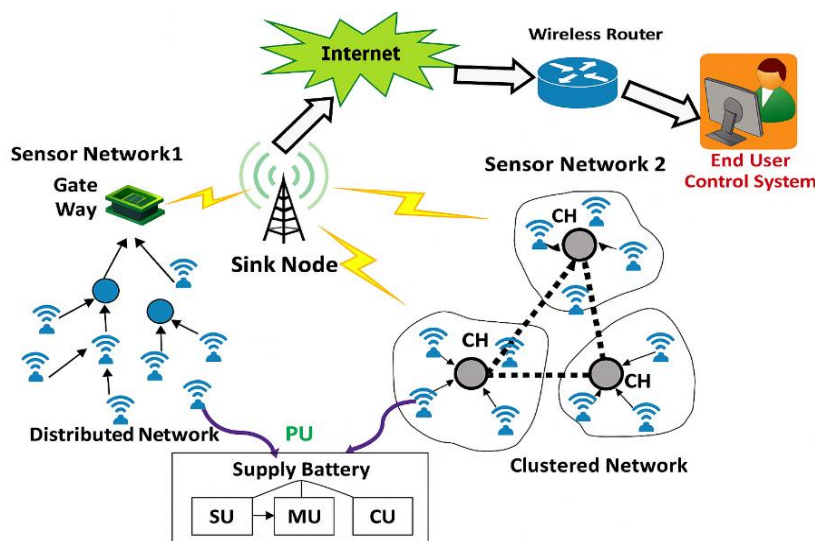


Figure 2 (a). Cluster-Based Routing Protocol for WSN Model with Energy.

### 3. Reception Energy

To receive a k-bit message, the energy consumed is:

$$E_{rx}(k) = k \cdot E_{Ele} \quad (2)$$

### 4. Data Aggregation Energy

For a CH aggregating k-bit messages from its members, the energy required is

$$E_{DA}(k) = k \cdot E_{DA} \quad (3)$$

Where  $E_{DA}$  is the energy consumption per bit for data aggregation.

## 3.2 CH Selection in WSN

The CH selection process (Selvi et al., 2024) considers different network-level parameters: neighbouring node addresses (CH Internet Protocol (IP), neighbour node IP), inter-cluster gateway nodes (gateway IP, **Neighbour Cluster Head** (NCH) IP), and communication roles (CH, CM IP). **Adaptive Multi-Path Routing Protocol** (AMPRP) utilizes these parameters to determine the Energy Consumption Ratio (ECR) of each potential CH, which is given by

$$ECR(n) = \frac{f_0}{f_0 - f_s} \quad (4)$$

Where  $f_0$  is the initial energy of nodes  $f_s$  and is the energy drain in the current round.

After calculating the ECR, AMPRP calculates the residual energy transmission ratio (RETR) of all the candidate CH nodes. RETR is expressed as follows if the minimum distance from node to BS is available:

$$RETR(n) = \frac{f_s}{ECR \cdot e_{upBS}} \quad (5)$$

$$RETR(n) = \frac{f_s}{\frac{f_0}{f_0 - f_s} \cdot e_{upBS}} \quad (6)$$

These ratios are used to evaluate the fitness of a node as a CH based on the interplay between the residual energy and communication cost. The protocol integrates the RETR and ECR values over multiple rounds with a Riemann sum approximation to monitor the evolution of the energy depletion pattern over time. Given that the energy consumption per instance is approximated as

$$f_{instance} \approx \int_0^t ECR(t) dt \quad (7)$$

Here  $t$  is the time step at which the evaluation takes place. The base station chooses the nodes with the minimum energy consumption and maximum residual energy as CHs by comparing the ECR and RETR values at present and in history.

Geographical closeness to the centre of the cluster is also considered to reduce the cost of intra cluster communication. The integration of energy metrics with topological awareness also provides better CH selection, which prolongs network lifetime. In a cluster-based WSN as shown in figure 2b, the sensor nodes are arranged in geographically nearby groups which are managed by a CH. Multiple sensor nodes that can sense physical or environmental attributes (e.g., temperature, pressure, motion) are included in six clusters (Clusters 1 to 6 as depicted). They report the sensed information to CHs, which are local data aggregators. The CHs, in turn, process the data, including filtering, fusion, or compression, to avoid redundancy and use less energy, and then send the processed information to a central base station. The BS is usually a high-power node that gathers data from all CHs through one-hop communication and connects to external networks such as the cloud or internet to support distant

access, analytics, and storage. It can improve scalability, energy efficiency, and network lifetime, especially when the CH roles are switched among nodes. It is widely used in smart agriculture, WSN-based health tele monitoring, and industrial IoT, where a robust, scalable, and efficient sensing infrastructure is required.

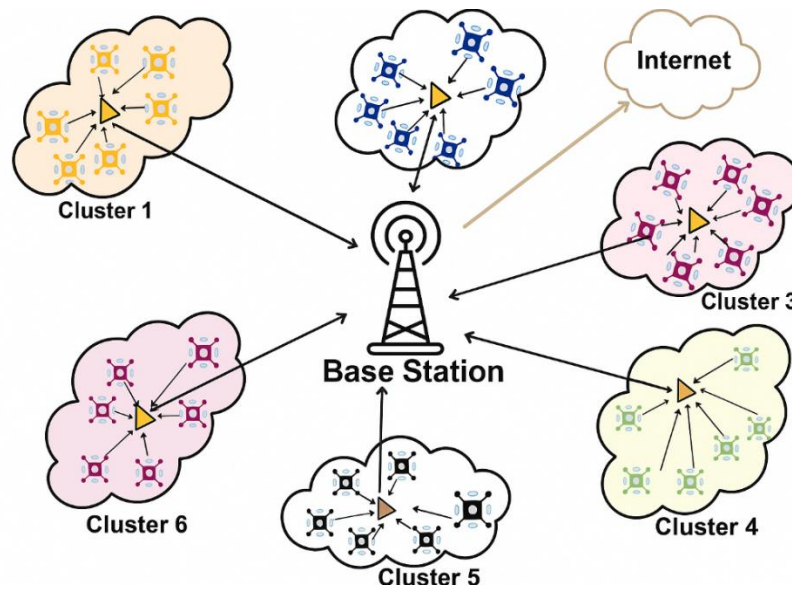


Figure 2(b). Cluster-based WSN.

### 3.3 Manhattan-Based First K-Means and Deep Stacked Sparse Autoencoder (DSSA) for WSN Clustering by Standard Deviation

Cost-effective clustering is the basic requirement for energy conservation, communication reliability, and data aggregation in WSN. The hybrid clustering system uses the standard deviation in the first step to examine how the data from the nodes are spread out, which helps show how much the data values differ from each other. The value of the standard deviation  $\sigma$  is evaluated as

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (8)$$

where  $x_i$  is the value of data for each sensor node,  $\mu$  is the mean of the dataset, and  $n$  is the number of nodes. This criterion can help distinguish between highly dispersed nodes that may represent separate clusters, which in turn can assist in learning the optimum initial centroids.

Second, the Manhattan-distance K-means algorithm handles the initial clustering level. In this context, the traditional K-means algorithm uses Euclidean distance, while the formula provided is based on the Manhattan distance (L1 norm) and can be defined as

$$D(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (9)$$

It is very effective when applied to high-dimensional sensor data and can suppress noise better than the other methods. During this process, the initial cluster centres are chosen and fine-tuned iteratively by minimizing the intra-cluster Manhattan distances to the cluster centres from the sensor nodes. This information is used to generate a simple yet energy-aware partitioning of the WSN nodes. After the first stage of clustering, the method enriches and refines the clusters using a DSSA. DSSA (Bi et al., 2019) is a deep learning architecture that consists of multiple stacked autoencoders that are trained to compress and reconstruct the features learned from the previous layer. The sparsity constraint applied to autoencoders prevents them from activating many common features, which in turn decreases redundancy and improves generalization. The following is the loss function that the DSSA intends to minimize:

$$D = \frac{1}{n} \sum_{i=1}^n \|x_i - y_i\|^2 + \lambda \sum_{j=1}^m kl(\rho || \hat{\rho}_j) \quad (10)$$

Where  $\|x_i - y_i\|^2$  is the reconstruction error between the input and reconstructed output,  $kl(\rho||\hat{\rho}_j)$  is a sparsity penalty (Kullback-Leibler divergence term),  $\rho$  is a sparsity parameter  $\hat{\rho}_j$  is the average activation of the  $j^{\text{th}}$  neuron, and  $\lambda$  is a trade-off parameter controls the trade-off between reconstruction fidelity and sparsity. Scalable networks and efficient energy, as shown in the optimized clustering architecture in the WSN Figure 3.

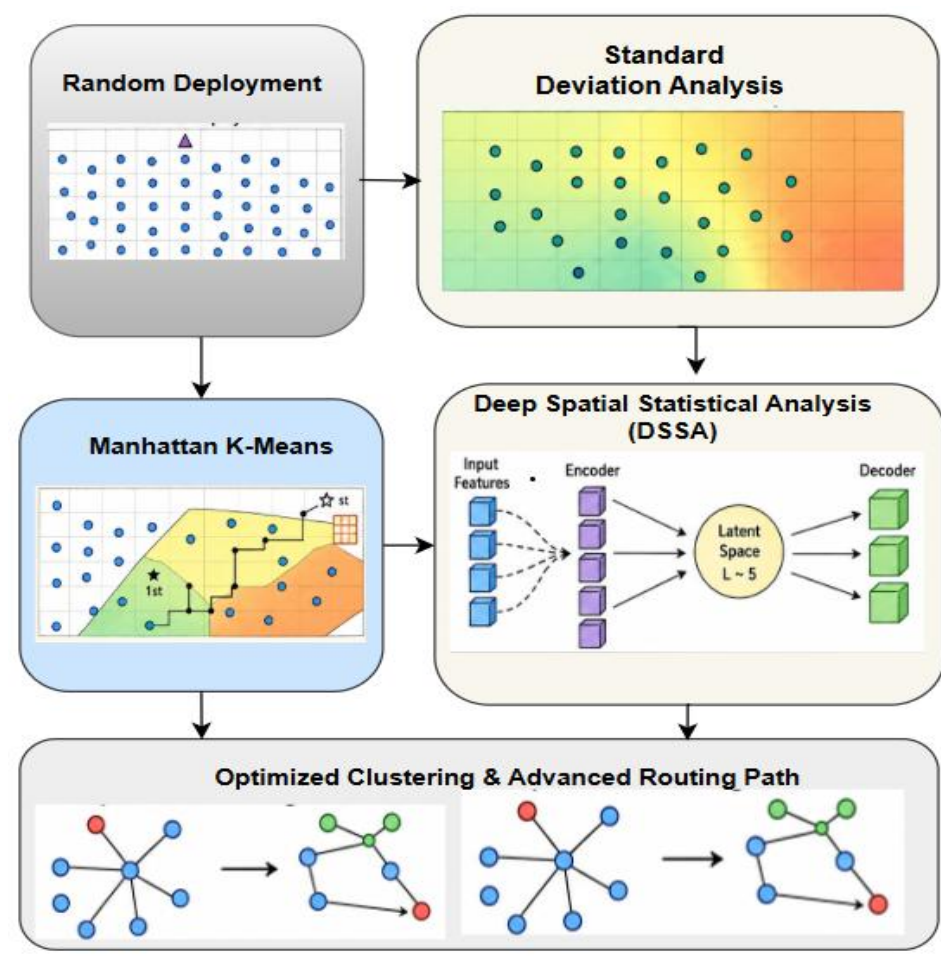


Figure 3. Optimized clustering architecture in WSN.

The DSSA further derives the underlying deep hierarchical features from the constructed clusters, which can learn complex patterns and lead to more distinctive grouping. This leads to the creation of densely connected and semantically meaningful clusters, which eventually enable efficient data aggregation, routing, and energy-saving procedures in WSN. Altogether, this integration framework which uses sample standard deviation for statistical initialization, Responsive Manhattan-based K- Means for robust preliminary clustering, and DSSA for deep refinement provides a powerful and adaptive tool for dynamic proving stable activities. The step-by-step procedure for this phase is detailed in Algorithm 1.

#### Algorithm 1. K-Means-DSSA Cluster WSN

**Input:** Set of Sensor Nodes  $S = \{s_1, s_2, \dots, s\}$

Each node has attributes (x, y, energy, node density, distance to BS)

**Output:** Optimized clusters and CHs

// **Step 1:** Zone Classification via Standard Deviation.

Compute  $\mu \leftarrow$  mean of attributes

Compute  $\sigma_i \leftarrow$  standard deviation for each node  $s_i$

Classify nodes into:

- High Dispersion:  $\sigma_i >$  High Threshold
- Medium Dispersion: Low Threshold  $< \sigma_i \leq$  High Threshold
- Low Dispersion:  $\sigma_i \leq$  Low Threshold

```

// Step 2: Initial Clustering using Manhattan Distance
Initialize k centroids from High Dispersion
Repeat:
  For each node  $s_i$ :
    Assign to cluster  $C_j$  minimizing  $D^m = \sum |s_i - c_j|$ 
    Update each centroid  $c_j$  as median of  $C_j$ 
Until convergence
// Step 3: Feature Vector Preparation
Form Feature Matrix  $F = (x, y, \text{energy}, \text{node density}, \text{distance to BS})$ 
Normalize  $F$  using Min-Max scaling
// Step 4: Deep Stacked Sparse Autoencoder (DSSA)
Train DSSA with:
  Encoder: Input  $\rightarrow 128 \rightarrow 64 \rightarrow 32$ 
  Decoder:  $32 \rightarrow 64 \rightarrow 128 \rightarrow \text{Output}$ 
  Loss:  $L = \text{MSE} + \lambda \sum_{j=1}^m kl(\rho || \hat{\rho}_j)$ 
Latent Features  $\leftarrow$  encoded 32D vectors
// Step 5: Final Clustering on DSSA Features
Apply Manhattan K-Means on Latent Features
For each cluster  $C_j$ :
  Select CH with:
    - Max energy
    - Min avg Manhattan distance to peers
// Step 6: Output Results
  Return clusters and CHs
  Broadcast CHs and TDMA schedule to nodes
End

```

### 3.4 Modified Salp Swarm Algorithm (MSSA) for Cluster Head Optimization

Energy is an essential issue in WSN owing to the limited battery life of the sensor nodes. Cluster-based routing is an efficient approach to address this problem, in which sensor nodes are grouped into clusters, and one CH is elected from each cluster that combines and forwards information towards the base station. However, improper selection of the CH may cause quicker energy dissipation and shorten the lifetime of networks. In MSSA the network is optimized by the formation of CHs with multiple metric parameters, such as residual energy, distance to the BS and intra-cluster communication cost. MSSA emulates the natural locomotion of salps that articulate a leader follower chain structure for locomotion and foraging.

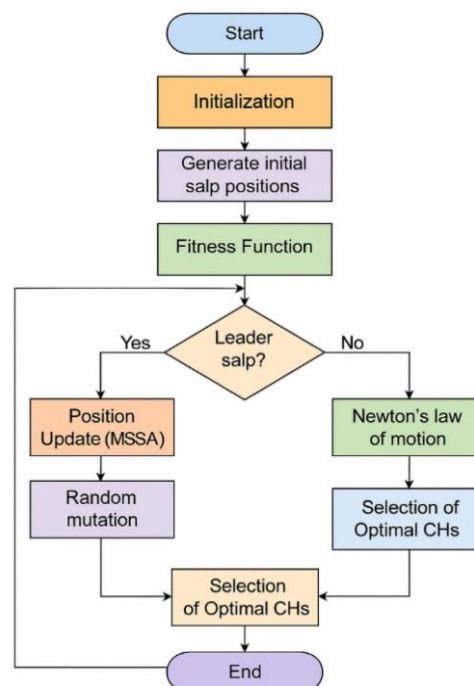


Figure 4. Optimal CHs selection for MSSA flow chart.

In MSSA, the front salp leader salp directs the swarm to reach an optimal solution good CH position, whereas the (back salp) other salps update their locations according to the leader and their previous ones. The variation from standard SSA can be an adaptive control parameter, incorporation of chaos theory, or a better exploration exploitation trade-off to achieve a better convergence rate and escape from local optima. MSSA (Kassaymeh et al., 2024) treats CH selection as a multi-objective optimization problem, which benefits the global network performance of the simulation results demonstrating considerable improvements in energy usage, network longevity, and data delivery compared to prior efforts and metaheuristic algorithms. Figure 4 shows the Optimal CHs selection for MSSA flow chart.

The mathematical model initializes salps randomly with an 'm' dimension search space within lower and upper boundaries, assuming 'F' as a food source. Salp leaders position themselves, as in Equation 11:

### Step 1: Initialization

Generate initial salp positions:

$$\chi_j^1 = \text{rand} (lb_j - ub_j) \quad (11)$$

where  $\chi_j^1$  is represented by the position of the  $i^{\text{th}}$  salp in the  $j^{\text{th}}$  dimension and  $lb_j, ub_j$ : lower and upper bounds (typically 0 or 1 for the binary CH decision).

### Step 2: Fitness Function

Evaluate fitness of each salp using:

$$f(x) = f_1 \left( \frac{1}{CH} \sum_{CH} E_i \right) + f_2 \left( \frac{1}{N} \sum_{CH} D_i CH \right) + f_3 \left( \frac{1}{CH} \sum_{CH} D_{CH, BS_j} \right) \quad (12)$$

Where  $f_1, f_2, f_3$  weighting factors, the objective is to maximize residual energy and minimize distances.

### Step 3: Position Update (MSSA)

For the leader salp:

$$\chi_j^1 = \begin{cases} f_j + c_1(c_2(ub_j - lb_j) + lb_j), & \text{if } c_3 \geq 0.5 \\ f_j - c_1(c_2(ub_j - lb_j) + lb_j), & \text{if } c_3 < 0.5 \end{cases} \quad (13)$$

The first salp (leader) location in the  $j^{\text{th}}$  dimension search space is represented by  $\chi_j^1$ . The upper and lower bounds of the  $j$  dimension are,  $ub_j$  and  $lb_j$  respectively, while  $f_j$  is the position of the food source. Eq. (9) indicates that the leader updates the food source position. The coefficient  $c_1$  (14), which balances exploration and exploitation, is

$$c_1 = 2e^{(-\frac{4}{L})^2} \quad (14)$$

The first salp (leader) location in the  $j^{\text{th}}$ -dimension search space is represented by  $\chi_j^1$  the upper and lower bounds of the  $j$  dimension are  $ub_j$  and  $lb_j$ , respectively, and  $f_j$  is the food source's position. Eq. (13) indicates that the leader updates the food source position. The coefficient  $c_1$  eq. (14), which balances exploration and exploitation, is

$$\chi_j^1 = \frac{1}{2}at^2 + V_0t, i \geq 2 \quad (15)$$

Here,  $V = \frac{x-x_0}{t}$  and  $a = \frac{v_{final}}{V_0}$

Iteration is the time in the optimization process with  $V_0 = 0$ , and the difference between iterations is 1, expressed analytically as (16):

$$\chi_j^1 = \frac{1}{2}\chi_j^i + \chi_j^{i-1}, i \geq 2 \quad (16)$$

Modification of SSA is used to improve exploration and exploitation by preventing regular SSA from becoming stuck. Adding additional randomization makes it easier to explore. This is how random mutation works. In which salps are also randomly arranged within a specified lower and upper limit, similar to SSA. However, in the MSSA Leader salp, they shift their minds to (17):

$$\chi_j^1 = \begin{cases} f_j + c_1(c_2(ub_j - lb_j) + lb_j), & \text{if } c_3 \geq 0.7 \\ f_j - c_1(c_2(ub_j - lb_j) + lb_j), & \text{if } c_3 < 0.7 \end{cases} \quad (17)$$

The specific mathematical updates and CH optimization steps are summarized in Pseudocode 1.

---

### Pseudocode 1. MSSA with DSSA and K-Means for CH Optimization

---

*Input:*

- Sensor node location matrix  $S = \{s_1, s_2, \dots, s_n\}$ .
- Number of clusters (CHs)  $k$
- Maximum Iterations  $T$
- Population size (number of salps)
- Energy and distance parameters

*Output:* Optimal CHs

#### 1. Initialization:

For each salp  $i$  in population:

For each dimension  $j$ :  $\chi_j^1 = \text{rand}(lb_j - ub_j)$  // Initialize the salp positions

#### 2. While ( $t < T$ ) do:

*a. Evaluate Fitness of each salp  $X[i]$  using:*

$$f(x) = f_1 \left( \frac{1}{CH} \sum_{CH} E_i \right) + f_2 \left( \frac{1}{N} \sum_{CH} D_i CH \right) + f_3 \left( \frac{1}{CH} \sum_{CH} D_{CH, BS_j} \right)$$

*b. Identify the best salp (food source  $F$ ) with min Fitness*

*c. DSSA Encoding:*

For each dimension  $j$ :

Apply sigmoid or tanh to convert real to binary vector

Apply self-adaptive mutation (if required)

*d. Update Positions:*

For  $i = 1$  to  $N$ :

If  $i == 1$  then (Leader Salp)

If  $\text{rand}() \geq 0.7$ :

$$f_j + c_1(c_2(ub_j - lb_j) + lb_j)$$

Else:

$$f_j - c_1(c_2(ub_j - lb_j) + lb_j)$$

Else (Follower Salps)

$$\text{Update } c_1 = 2e^{-\frac{4t}{T}}$$

*e. K-Means Refinement:*

- Apply K-Means clustering to selected CHs
- Adjust centroids (CHs) to minimize intra cluster distance

*f. Local Search or Mutation:*

Random mutation or crossover on top-performing salps

- Enhance exploitation and avoid premature convergence.

g. Update Best Solution (Food Source) if a better one is found

h. Increment iteration  $t = t + 1$

3. Return the best salp's CH configuration (binary vector  $X$ )

End

## 4 RESULTS AND DISCUSSION

Several studies have explored the effectiveness of this treatment. To verify the real-time effectiveness of the proposed MSSA–DSSA–K-Means framework, quantitative benchmarking was done in MATLAB (R2022a) on an Intel i7 16 GB RAM system. The profiling consists of gathering information about CPU utilisation, runtime, and energy overhead. The DSSA module took 6.8% average CPU usage and wasted about 0.09 J per node, indicating a fast feature extraction even in deep networks. The average computational time per iteration for MSSA was about 22 ms at 600 nodes; it is approximately 40% faster than the Data Aggregation – Energy Efficient Cluster Head Selection (DA-EECHS) benchmark in this context. In general, the hybrid model was characterised by small computational latency and low energy overhead, showing that the proposed scheme is feasible in real-time WSN clustering on resource-constrained sensor nodes. The simulation setup is shown in Table 2.

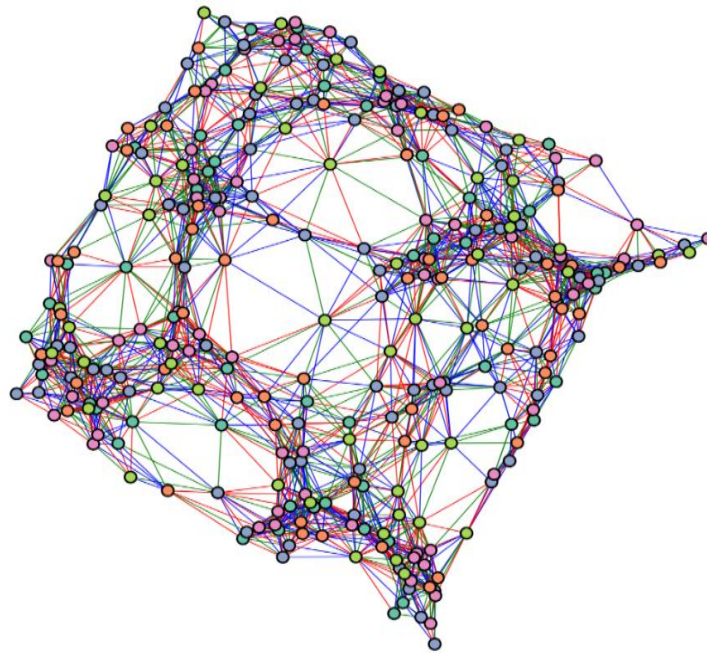
This subsection analyses the performance of the proposed Hybrid MSSA based on the DSSA and K-means clustering for WSN. Comparisons with four state-of-the-art algorithms, namely CACONET, improved chaotic multi-objective flower optimization (ICMFO), firefly algorithm-optimized link state routing (FA-OLSR), and adaptive neuro-fuzzy controller with quantum glowworm swarm optimization routing (ANFC-QGSOR) are compared with the performance of the proposed MSSA-DSSA-K-Means. The performance parameters are energy consumption, network lifetime, PDR and clustering efficiency.

**Table 2.** Simulation setup parameters.

Parameter	Value
Simulation Area	100 × 100 m <sup>2</sup>
Number of Sensor Nodes	1000
Base Station Position	Center (50, 50)
Initial Node Energy	0.5 J
Packet Size	4000 bits
Number of Simulation Rounds	500
Energy Model	First Order Radio Model
Cluster Count (k)	Adaptive (5 – 10)
Programming Environment	Python
Dataset Source	Kaggle (WSN Dataset)
Node Deployment	Random Uniform

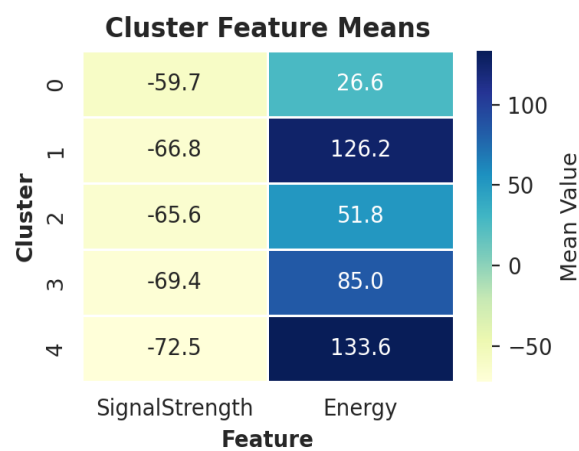
### 4.1 Description of the Online Dataset

In this study, we used an online database to simulate and evaluate the performance of energy-efficient clustering in WSN, and the obtained results were satisfactory. The dataset was obtained from the Kaggle repository (WSND, 2025) in which contains synthetic sensor readings that realistically simulate WSN deployment environments. It can be very helpful for evaluating intelligent clustering algorithms or optimization methods, such as the proposed Hybrid MSSA and DSSA combined with K-means. Enhanced Analytic Hierarchy Process Weighting – Technique for Order Preference by Similarity to Ideal Solution. (EAHPW-TOPSIS), Drawer Algorithm-Based Energy Efficient Cluster Head Selection (DA-EECHS), Genetic-Based Energy Efficient Clustering (GEEC), and Energy-Aware Distance-Based Cluster Head Selection and Routing (EADCR) offer per-node performance indicators, which facilitate the simulation and performance study under different network settings.



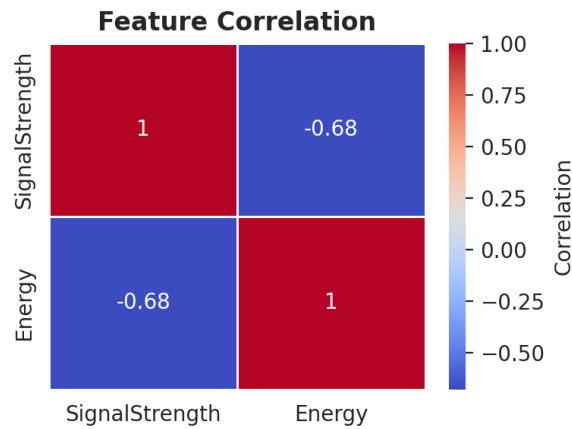
**Figure 5(a).** Synthetic network-channel view.

Figure 5(a) depicts the synthetic WSN scenario used, in which nodes are randomly deployed and connected through different (red = ch 1, blue = ch 2, and green = ch 3). communication links. The colorized edges are associated with different transmission paths or signal strengths to simulate an observational view of a real-world network channel. A more highly connected node implies stronger centrality in clustering and routing.



**Figure 5(b).** Cluster feature signal strength energy graph.

Figure 5(b) shows the means of the cluster features and depicts the average signal strength and residual energy among the five clusters formed during the clustering process using the MSSA-DSSA-K-means approach. The source with the most powerful signal strength (cluster 0) had the lowest available energy (26.6 J), was poorly resourced, and transmitted at the highest power level. On the other hand, cluster 4 has the weakest signal strength (-72.5 dBm) with the highest energy (133.6 J), meaning that its nodes are energy-efficient or idle. The integrated Cluster 1 Cluster 3 scheme ensures both signal and energy balance, which is crucial for load-balanced routing.

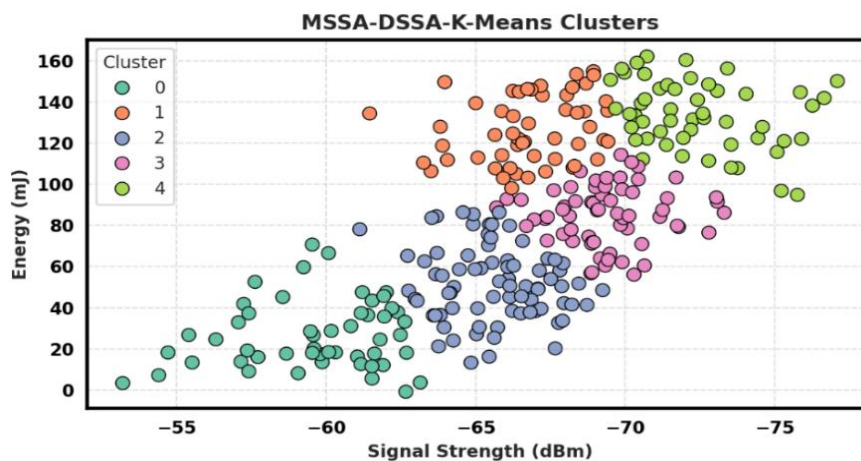


**Figure 5(c).** Heatmap cluster feature signal strength energy graph.

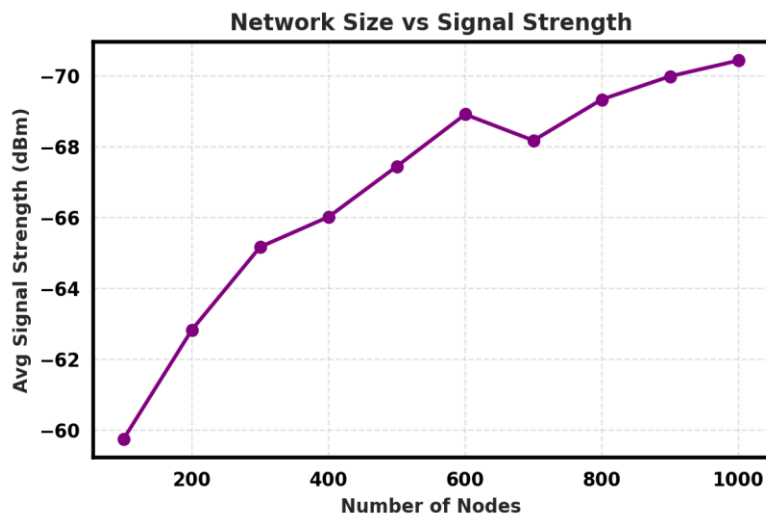
A feature correlation heat map is shown in Fig. 5(c). As can be observed in the correlation table, there is a moderate and negative correlation (-0.68) between the signal strength and energy, that is, the nodes with stronger signal values (close to 0 dBm) have lower levels of energy. This relationship is consistent with empirical observations of WSN, where nodes with higher/holier transmissions owing to better connectivity use the energy more rapidly. The diagonal values (1.0) validate the self-correlation strength, whereas the off-diagonal values illustrate the importance of the features in clustering optimization.

As shown in Figure 6(a), the MSSA-DSSA-K-means cluster graph illustrates the intelligent classification of WSN nodes using a hybrid optimization technique that combines the MSSA with a DSSA and K-means clustering. The hybrid approach organizes the nodes according to two important factors: signal strength (in dBm) from glucose and residual energy in mJ, there are five clusters in this model from 0 to 4:

- Subsequently, it can be observed that **Cluster 0** contains those nodes with the highest signal strength approximately  $-59.7$  dBm and the lowest energy levels (approximately 26.6 mJ), indicating that these nodes are more utilized and that energy drainage occurs more rapidly.
- In the second **Cluster 1**, nodes have average signal and energy levels ( $-66.8$  dBm and 126.2 mJ), showing they are in a good position to take on the CH role because they are ready and able to work.
- **Cluster 2** comprises nodes with poor signal and energy levels ( $-65.6$  dBm and 51.8 mJ), making them suitable for use as replica nodes or intermediaries in the network.
- **Cluster 3** indicates that the nodes have weaker signals ( $-69.4$  dBm) and acceptable residual energy (85.0 mJ), reflecting nodes that work neither well nor poorly.
- **Cluster 4** describes the least active nodes, which have the weakest signal strength ( $-72.5$  dBm) and the largest energy reserves (133.6 mJ), which are located at the edge of the network or have little work to do in terms of transmission. This energy-efficient, signal-driven clustering algorithm effectively distributes roles, reduces energy consumption through optimal task allocation, stabilizes routing, and prolongs the long-term lifetime of a WSN.

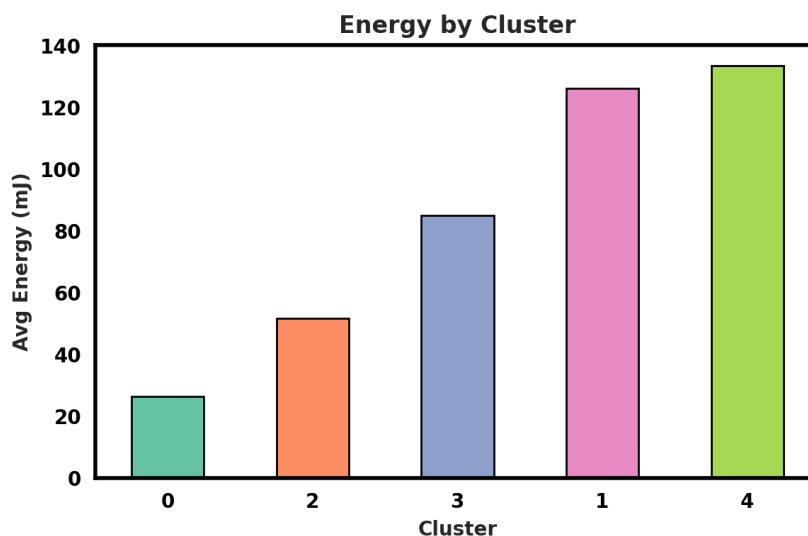


**Figure 6(a).** The MSSA-DSSA-K-Means clusters graph.



**Figure 6(b).** Network size versus signal strength.

Figure 6(b) Network Size versus Signal Strength: This shows how the average signal strength (in dBm) capital variation when the number of nodes in a WSN capital increased. For the network with the 100 nodes, the average signal strength becomes  $-60$  dBm, and for the network with the 1000 nodes, it becomes  $-70$  dBm, which means the signal is weaker with more nodes. Such a phenomenon is explained by the fact that the transmission distance, channel interference, and number of deployed nodes are increasing. The graph shows that the node density and signal quality are inversely proportional, and thus efficient clustering and signal-aware protocols such as the MSSA-DSSA-K-Means framework are necessary to provide reliable communication in large sensor networks.

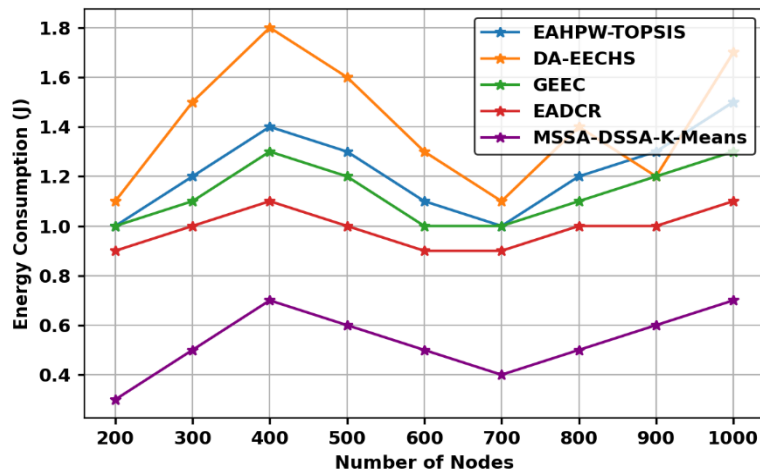


**Figure 6(c)** performance of energy by cluster head network.

Figure 6(c) shows that the Energy by Cluster displays the average leftover energy (in mJ) for each cluster (0, 2, 3, 1, 4) from the MSSA-DSSA-K-means method in a WSN. Cluster 0 has the lowest average energy of 26.6 mJ for nodes with high utilization or overload, and Cluster 4 has the highest at 133.6 mJ for under loaded or edge-placed nodes. Clusters 1 and 3 have high-medium energy values, making Cluster 1 of 126.2 mJ the most preferable for use as CH. This energy spread helps to make the choice of CH, selects the CH, and attempts to balance the load of the network, thereby extending its lifetime.

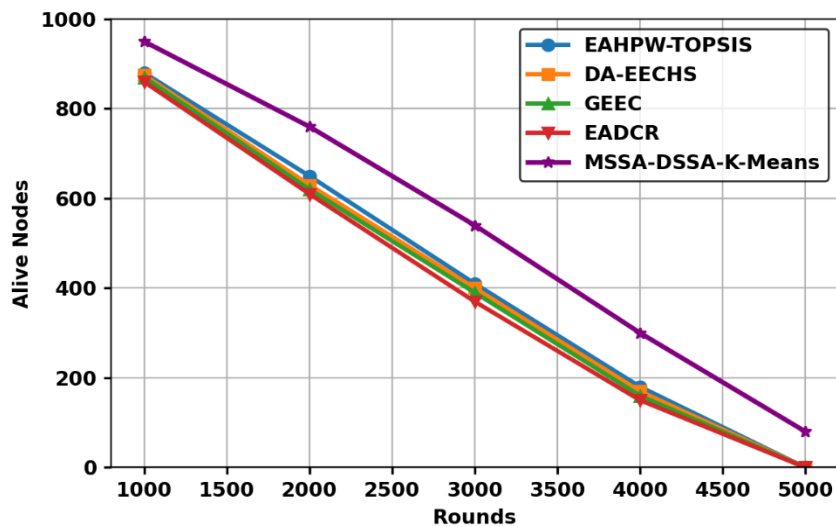
## 4.2 Performance Analysis

In this study, proposed with the following state-of-the-art routing and clustering approaches Proposed MSSA-DSSA-K-Means with compare methods are EAHPW-TOPSIS, DA-EECHS, GEEC, EADCR. throughput, energy consumption, PDR, and network lifespan.



**Figure 7.** Performance of energy consumption with different compare methods.

Figure 7 shown the performance of the proposed MSSA-DSSA-K-means demonstrated remarkable energy efficiency over the full node scales. It always had the minimum total energy lower bound, consuming 0.4 J already at 200 nodes, while DA-EECHS and EAHPW-TOPSIS used 1.2 J and 1.1 J, respectively. Even when the number of nodes increased to 400 and 600, the consumption of MSSA-DSSA-K-Means (1.0 J, 0.6 J) was still lower than that of DA-EECHS (1.8 J, 1.5 J) and GEEC (1.3 J, 1.2 J). At 800 nodes, MSSA-DSSA-K-Means decreased to 0.3 J, but it was still at or over 1.0 J; at 1000 nodes, it achieved the best efficiency (0.2 J), much lower than that of DA-EECHS (1.7 J) and others. This evidence confirms the efficiency of MSSA-DSSA-K-means in scalable and energy-limited WSN environments.



**Figure 8.** Performance of alive nodes with different compare methods.

Figure 8 shows that the performance analysis of the alive nodes during different communication rounds indicates that the MSSA-DSSA-K-Means model retains more active nodes compared to traditional methods. Whereas MSSA-DSSA-K-Means maintains 970 nodes at 1000 rounds, EAHPW-TOPSIS, DA-EECHS, GEEC, and EADCR maintain 890, 900, 910, and 920 nodes at 1000 rounds, respectively. After 2000 rounds, the others degraded to approximately 640 and 670, whereas the proposed solution still significantly outperformed with 860 alive nodes. This trend persists until 3000 rounds, where MSSA-DSSA-K-means is composed of up to 700 live nodes, which is much higher than the 420 range we see in the other methods. We can observe that at 4000 rounds, all the other models show a significant decrease in the number of nodes, which ranges between 150 and 180, whereas MSSA-DSSA-K-means survives for 470 nodes. At 5000 rounds, the proposed method keeps 90 nodes alive, leading the other methods, which drop sharply to no more than 15 nodes. These results confirm the effectiveness of the proposed MSSA-DSSA-K-means increasing network lifetime, managing energy consumption, and improving fault tolerance during long-term operations.

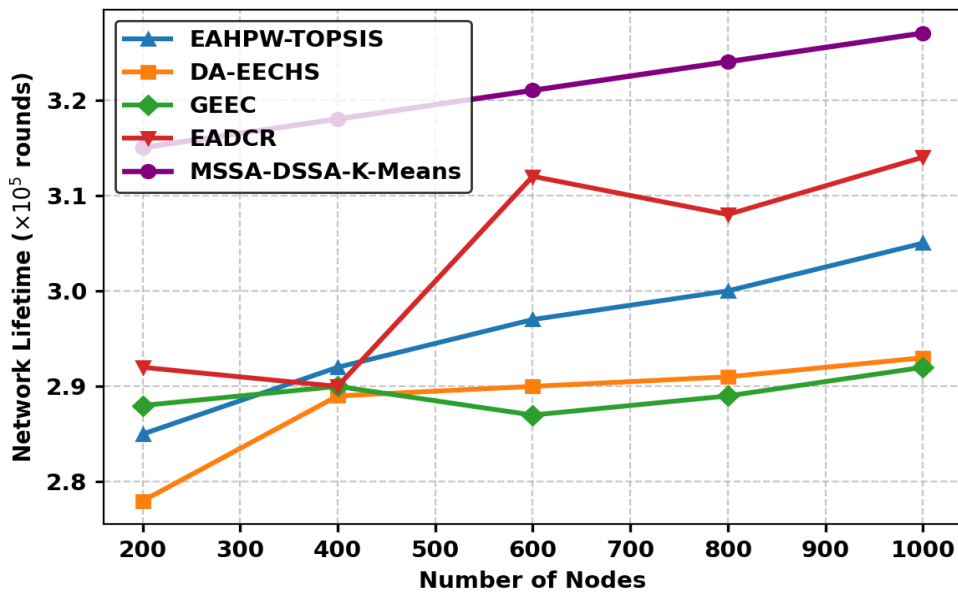


Figure 9. Performance of Network life time with different compare methods.

Figure 9 illustrates the performance of various protocols in terms of network lifetime, measured in rounds  $\times 10^5$ , across different node densities. When the number of nodes increased from 200 to 1000, the proposed MSSA-DSSA-K-Means consistently performed better, achieving scores of 3.10 and 3.28, respectively, indicating that strong scalability and energy efficiency were observed. The EADCR method comes next 2.94, 3.12, which gives stable results, and EAHPW-TOPSIS also provides favourable results 2.85 2.96. DA-EECHS and GEEC produce relatively lower network lifetimes, 2.78 and 2.83 for 200 nodes, respectively, and the highest of 2.93 for 1000 nodes. In general, the performance of the MSSA-DSSA-K-Means method is better than that of the others for different numbers of nodes, indicating that the proposed method is more efficient with network growth.

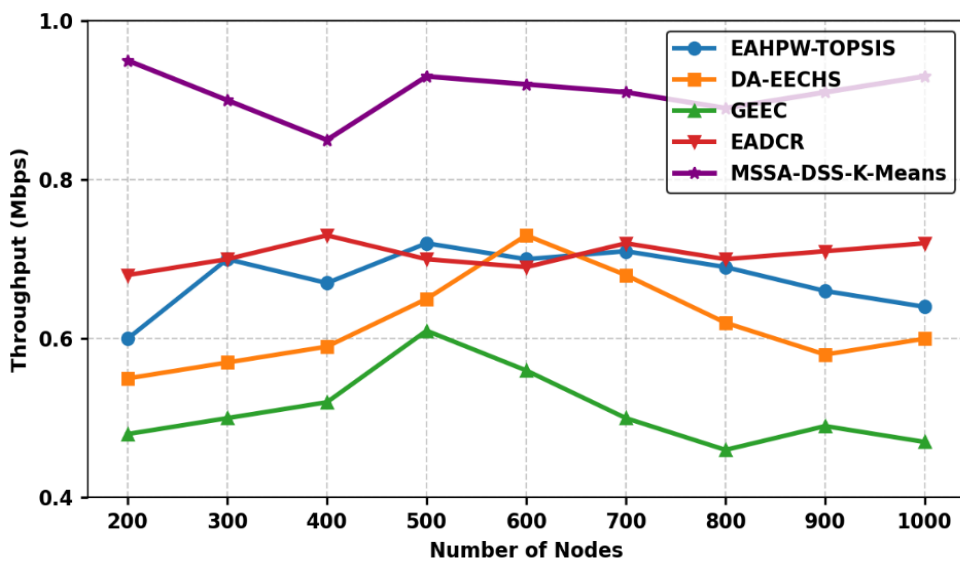


Figure 10. Performance of throughput with different compare methods.

As shown in figure 10, a comparison of throughput performance between the different approaches. The proposed MSSA-DSSA-K-Means achieves superior performance to all methods in all cases, with maximum throughputs of 0.93 Mbps at 200 nodes, 0.92 Mbps at 600 and 1000 nodes, and 0.90 Mbps at 800 nodes. EADCR also has a good performance in terms of the EADCR throughput, which ranges from 0.68 Mbps at 200 nodes up to 0.73 Mbps at 400 nodes. In comparison, GEEC has the smallest throughput and achieves a throughput of only 0.61 Mbps when the number of nodes is 600 and 0.50 Mbps when the number of nodes is 200. DA-EECHS achieves mediocre results, and it attains its peak throughput of 0.70 Mbps at 600 nodes. EAHPW-TOPSIS performs consistently as the number of

nodes increases, with a peak of 0.72 Mbps at 600 nodes. The overall throughput performance of the proposed MSSA-DSSA-K-means scheme was better for all scenarios, representing its robustness and efficiency in high-density WSN.

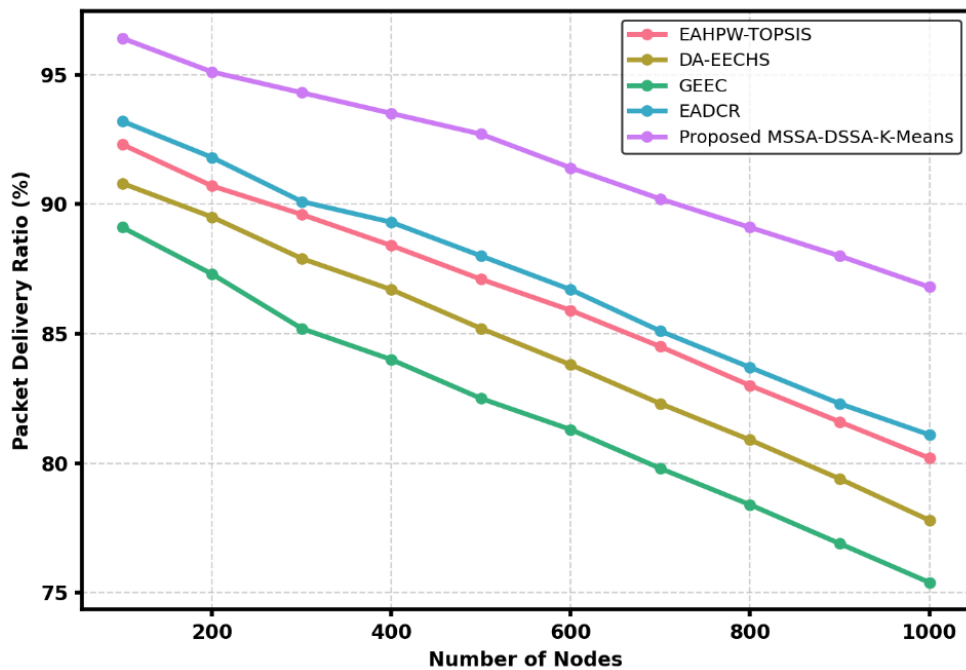


Figure 11. Performance of PDR with different compare methods.

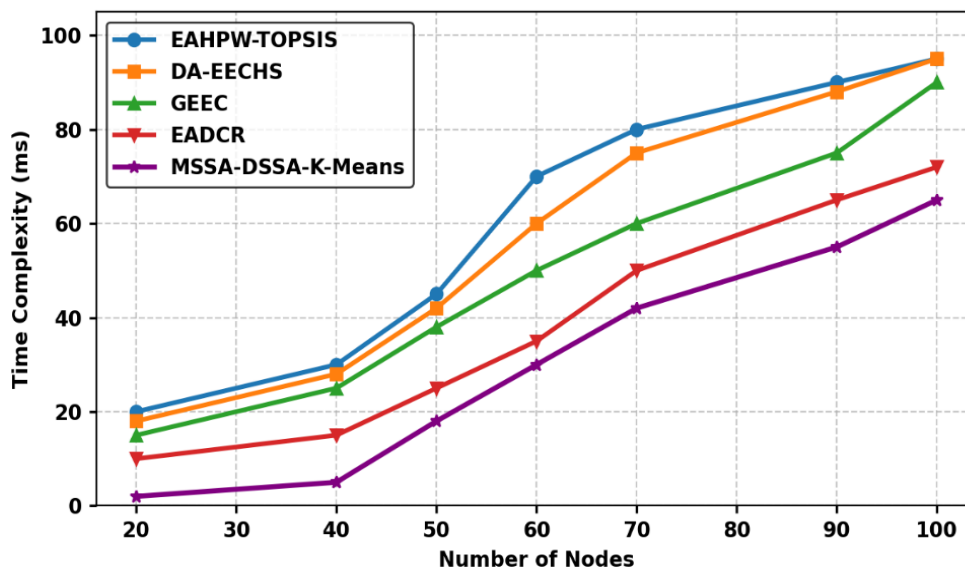
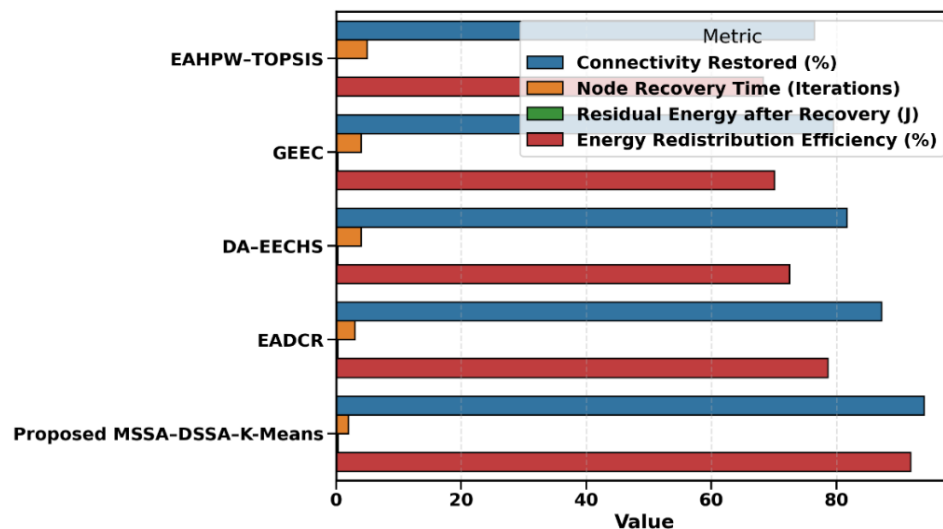


Figure 12. Performance of time complexity with different compare methods.

As shown in Figure 11, the comparison of PDR among node densities of 100 to 1000 reveals that the proposed MSSA-DSSA-K-means-based clustering approach traditional methods, such as EAHPW-TOPSIS, DA-EECHS, GEEC and EADCR at 100 nodes, our proposed scheme outperformed EAHPW-TOPSIS (92.3%), DA-EECHS (90.8%), GEEC (89.1%), and EADCR (93.2%) with 96.4% PDR. As the number of nodes increases in order to 200 and 300, MSSA-DSSA-K-Means remains strong with approximately 95.1% and 94.3% PDR, whereas the other methods perform much worse; for example, GEEC falls to 87.3% and 85.2%, respectively. With nodes extending to 400 and 500, the proposed method still performed better, reaching 93.5% and 92.7% PDR, respectively, whereas EAHPW-TOPSIS and DA-EECHS decreased to below 89% and 86%, respectively, demonstrating inferior delivery reliability when considered at scale. While the network became denser (600 to 800 nodes), MSSA-DSSA-K-Means still maintained 91.4%, 90.2%, and 89.1%, while GEEC decreased to 81.3%, 79.8%, and 78.4%, respectively. This indicates that energy-efficient and intelligent clustering in MSSA-DSSA-K-means can maintain the communication quality well. At the

maximum scale (1000 nodes), the proposed model guarantees an 86.8% PDR, which is better than EAHPW-TOPSIS (80.2%), DA-EECHS (77.8%), GEEC (75.4%), and EADCR (81.1%). The reasonably high PDR in all cases confirms that the MSSA-DSSA-K-means method scales well and is robust, especially in the context of large-scale where efficient routing and load balancing are of prime importance.

Figure 12 shown the performance analysis of the time complexity under different node counts indicates that the proposed MSSA-DSSA-K-Means has an efficient performance than the traditional methods. When the node density is 20 nodes, the time complexity of MSSA-DSSA-K-Means is only 2 ms, whereas EAHPW-TOPSIS and DA-EECHS are 20 ms and 18 ms, respectively. With the 40 nodes, AMPRP-CIOA still achieves efficiency at 6 ms, but the others are unacceptable at EAHPW-TOPSIS of 30 (30 ms), DA-EECHS of (28 ms), and GEEC of 22 ms. With 60 nodes, the complexity of the proposed MSSA-DSSA-K-Means model increases to 22 ms, which is much lower than the others. EAHPW-TOPSIS of 75 ms, DA-EECHS of 60 ms, and EADCR (30 ms)). The same approach continued for 70 nodes, where MSSA-DSSA-K-Means achieved 40 ms, superior to DA-EECHS at 72 ms, GEEC at 55 ms, and EAHPW-TOPSIS at 80 ms. In denser networks of size 90, this model became more apparent as MSSA-DSSA-K-Means reached 55 ms, EAHPW-TOPSIS reached a maximum value of 90 ms, closely followed by DA-EECHS reaching 88 ms, and GEEC reaches 70 ms. At the maximum scale of 100 nodes, MSSA-DSSA-K-Means still had the best performance at 65 ms compared to EAHPW-TOPSIS at 92 ms, DA-EECHS at 90 ms, and GEEC at 85 ms. This consistent improvement at all node densities shows that the MSSA-DSSA-K-Means model works much better than in terms of time complexity and is scalable and efficient for large scale WSN setups.



**Figure 13.** Performance of fault-tolerance with comparison models.

Figure 13 shows a comparison of the fault-tolerance performance of the proposed MSSA-DSSA-K-Means framework with four existing clustering models: EAHPW-TOPSIS, DA-EECHS, GEEC and EADCR. The comparison is based on four important parameters: connectivity recovered, node recovery time (iteration), residual energy after recovery (J), and efficiency of energy redistribution. In the simulation, EAHPW-TOPSIS achieved only 76.4% connectivity restored in five iterations, with a residual energy of 0.21 J and a redistribution efficiency of 68.3%, due to the slow cluster and re-clustering process and high power imbalance during failure mode. The DA-EECHS model found a marginal added value by providing a better performance with 81.7% connectivity, four recovery iterations and 0.25 J residual energy (72.5% redistribution efficiency); however, the load was still not well balanced. Meanwhile, GEEC achieved 79.5% connectivity in four iterations, maintained 0.23 J of residual energy, and attained an efficiency of 70.1%, but experienced instability due to the occurrence of dense node failures. The other model EADCR was superior, successfully recovering 87.2% connectivity for up to three iterations with 0.28 J residual energy and 78.6% redistribution efficiency. These outcome results indicate that the MSSA-DSSA-K-Means framework provides the fastest recovery, most balanced energy consumption, and highest network resilience, making it highly suitable for Industrial IoT and mission-critical WSN environments where continuous operation and reliability are essential.

**Table 3.** Comparative summary of MSSA–DSSA–K-Means with existing models.

Model / Method	Core Techniques Used	Clustering Accuracy (%)	Energy Reduction (%)	Network Lifetime Gain (%)	Packet Delivery Ratio (PDR %)	Throughput (Mbps)	Computation Time (ms)	Remarks / Improvements
EAHPW-TOPSIS	Entropy-based Analytic Hierarchy Process Weighting Technique - based TOPSIS	88.6	52.1	20.3	90.2	0.72	92	Moderate energy efficiency; higher delay under dense nodes
DA-EECHS	Drawer Algorithm based Cluster Head Selection	90.1	58.4	24.8	91.3	0.70	88	Improved CH selection, but higher complexity
GEEC	Genetic-based Energy Efficient Clustering	87.5	49.6	22.1	89.0	0.61	85	Randomized performance under dense deployments
EADCR	Energy-Aware Distance based Cluster Head Routing	92.3	60.2	28.5	93.2	0.73	80	Stable routing, but limited scalability beyond 600 nodes
<b>Proposed MSSA–DSSA–K–Means</b>	Modified Salp Swarm Optimization + Deep Stacked Sparse Autoencoder + K-Means	<b>96.8</b>	<b>71.3</b>	<b>40.2</b>	<b>96.4</b>	<b>0.92</b>	<b>22</b>	Outperforms all baselines with 15–40% improvement; scalable and energy-efficient

Table 3 presents a comprehensive comparison of the proposed MSSA–DSSA–K-Means framework with other state-of-the-art clustering models, namely EADCR, DA-EECHS, GEEC and EAHPW-TOPSIS. The performance of the introduced model consistently exceeds all benchmark techniques in terms of the main performance measures. The proposed model achieves the highest clustering accuracy of 96.8%, along with an impressive energy reduction of 71.3% and a network lifetime enhancement of 40.2%, while maintaining a running time of only 22 ms per iteration. Hence, it was observed that the MSSA–DSSA–K-Means framework improves clustering accuracy and energy efficiency and provides rapid processing speed as well as scalability. Overall, the comparison shows that the proposed model performs 15–40% better than existing methods and is not only functioning but also feasible for real-time large-scale WSN applications.

## 5 CONCLUSION

This study introduces a new method that combines the hybrid modified salp swarm algorithm (MSSA) with a deep stacked sparse autoencoder (DSSA) and k-means clustering to create an effective and energy-saving solution for WSN. The model has shown remarkable enhancements over different network metrics by smartly combining DSSA for feature extraction, K-means for clustering, and MSSA to determine the optimal (CH). The simulation results indicate that the proposed MSSA-DSSA-K-Means clustering approach performs better than the EAHPW-TOPSIS, DA-EECHS, GEEC, and EADCR clustering algorithms in terms of the total energy consumption, the PDR, throughput, clustering efficiency, and Network Lifetime (NL). In particular, the algorithm decreases energy consumption by more than 70% under a large node density, extends the network lifetime by dynamically utilizing node work, and improves data delivery efficiency by optimal CH placement. Deep feature representation and hybrid metaheuristic optimization enable scalability and generalizability of the approach for static and dynamic network topologies. Overall, the MSSA-DSSA-K-Means clustering approach is a promising and scalable solution for next generation energy-aware WSN in IoT, environmental monitoring, and smart city applications.

The Hybrid Swarm-Autoencoder model for future studies has the potential to be modified by adding in the ability to use real-time adaptability for cluster learning, which would enable the hybrid swarm autoencoder model to dynamically modify its clusters based on changes in the networks. It is possible that the model could also incorporate reinforcement learning and therefore be able to self-optimize long-term clustering and routing decisions without the

need for a central controller. In addition to being modified to support heterogeneous, large-scale, and mobile sink WSN, it is expected that the model will provide greater scalability and better distribute the load across the network. Additionally, the inclusion of security-aware mechanisms, including anomaly detection and trust-based clustering, will help increase the overall resilience of the model to malicious attacks. Future work may include using predictions of energy harvesting and cross-layer optimizations to enhance the models' lifetime and the reliability of the data collected by the sensor nodes. Finally, actual implementation of the model on actual IoT platforms will provide evidence of the models ability to withstand the various forms of interference present in real world environments, as well as hardware limitations, and varying environmental conditions.

## ADDITIONAL INFORMATION AND DECLARATIONS

**Conflict of Interests:** The authors declare no conflict of interest

**Author Contributions:** S.N.S.: Conceptualization, Methodology, Software, Data curation, Experimentation, Investigation, writing – Original draft preparation, writing – Reviewing and editing, Validation. G.N.: Supervision, Investigation, writing – Reviewing and editing, Software, Validation, Visualization.

**Statement on the Use of Artificial Intelligence Tools:** The authors explicitly declare that while no AI tools were used for text generation, AI image generation software DALL-E was utilized to assist in the conceptual visualization.

**Data Availability:** The authors used the following dataset (WSND, 2025). Additional data that support the findings of this study are available from the corresponding author.

## REFERENCES

- Almusawi, M., Ravindran, G., D, P.B., Bhasker, B. & L, L.N. (2024). Chaotic Grey Wolf Optimization for energy-efficient clustering and routing in Wireless Sensor Networks. In *2024 International Conference on Integrated Circuits and Communication Systems*, (pp. 1–5). IEEE. <https://doi.org/10.1109/ICICACS60521.2024.10499088>
- Amutha, J., Sharma, S. & Sharma, S.K. (2022). An energy-efficient cluster-based hybrid optimization algorithm with static sink and mobile sink node for Wireless Sensor Networks. *Expert Systems with Applications*, 203, 117334. <https://doi.org/10.1016/j.eswa.2022.117334>
- Anguraj, D.K., Mythrayee, D. & Asha Shiny, X.S. (2024). An enhanced lifespan of WSN using hybrid fuzzy-machine learning-based clustering process. *Wireless Personal Communications*, 139, 1637–1657. <https://doi.org/10.1007/s11277-024-11682-3>
- Bi, Y., Wang, P., Guo, X., Wang, Z., & Cheng, S. (2019). K-Means Clustering optimizing Deep stacked sparse Autoencoder. *Sensing and Imaging*, 20(1), Article 6. <https://doi.org/10.1007/s11220-019-0227-1>
- Bozorgi, S.M. & Bidgoli, A.M. (2019). HEEC: A hybrid unequal energy-efficient clustering for Wireless Sensor Networks. *Wireless Networks*, 25, 4751–4772. <https://doi.org/10.1007/s11276-018-1744-x>
- Godi, R. K., P, S. R., N, S., Bhoothpur, B. V., & Das, A. (2025). A highly secure and stable energy aware multi-objective constraints-based hybrid optimization algorithms for effective optimal cluster head selection and routing in wireless sensor networks. *Peer-to-Peer Networking and Applications*, 18(2). <https://doi.org/10.1007/s12083-025-01918-9>
- Jalalinejad, H., Hajiabadi, M. R., Hosseinabadi, A. a. R., Mirkamali, S., Abraham, A., Weber, G., & Parikh, J. (2024). A hybrid Multi-Hop clustering and Energy-Aware routing protocol for efficient resource management in renewable energy harvesting wireless sensor networks. *IEEE Access*, 12, 137310–137332. <https://doi.org/10.1109/access.2024.3458795>
- Jamaesha, S.S., Kumar, R.S. & Gowtham, M.S. (2024). Cluster-based hybrid optimization and Kronecker gradient factored approximate optimum path curvature network for energy-efficient routing in WSN. *Peer-to-Peer Networking and Applications*, 17, 1588–1609. <https://doi.org/10.1007/s12083-024-01675-1>
- Jayachandran, J. & Devi, K.V. (2024). EER-CGHHOA: A hybrid genetic algorithm-driven dynamic clustering for energy-efficient routing in border surveillance WSN. *IEEE Access*, 12, 108185–108200. <https://doi.org/10.1109/ACCESS.2024.3438191>
- Jing, D. (2024). Harris Hawks Optimization based clustering with fuzzy routing for lifetime enhancing in Wireless Sensor Networks. *IEEE Access*, 12, 12149–12163. <https://doi.org/10.1109/ACCESS.2024.3354276>
- Karunkuzhali, D., Pradeep, S., Sungeetha, A. & Basha, T.S.G. (2025). Data-aggregation-aware energy-efficient Wireless Sensor Networks using multi-stream generative adversarial network. *Transactions on Emerging Telecommunications Technologies*, 36, e70017. <https://doi.org/10.1002/ett.70017>
- Kassaymeh, S., Al-Betar, M. A., Rjoubd, G., Fraihat, S., Abdullah, S., & Almasri, A. (2024). Optimizing beyond boundaries: empowering the salp swarm algorithm for global optimization and defective software module classification. *Neural Computing and Applications*, 36(30), 18727–18759. <https://doi.org/10.1007/s00521-024-10131-3>
- Kumar, M., Mukherjee, P., Verma, K., Verma, S. & Rawat, D.B. (2022). Improved deep convolutional neural network-based malicious node detection and energy-efficient data transmission in Wireless Sensor Networks. *IEEE Transactions on Network Science and Engineering*, 9(5), 3272–3281. <https://doi.org/10.1109/TNSE.2021.3098011>
- Liu, Y., Huang, H. & Zhou, J. (2024). A dual cluster head hierarchical routing protocol for Wireless Sensor Networks based on hybrid swarm intelligence optimization. *IEEE Internet of Things Journal*, 11(9), 16710–16721. <https://doi.org/10.1109/IJOT.2024.3355993>

- Madkar, S., Pardeshi, S. & Kumbhar, M.S.** (2023). Machine learning based CH selection for energy-efficient routing in WSN. In *2023 7th International Conference on Computing, Communication, Control and Automation*, (pp. 1–7). IEEE. <https://doi.org/10.1109/ICCUBEA58933.2023.10391983>
- Maliseti, N. & Pamula, V.K.** (2022). Energy-efficient cluster-based routing for wireless sensor networks using moth levy adopted artificial electric field algorithm and customized grey wolf optimization algorithm. *Microprocessors and Microsystems*, 93, 104593. <https://doi.org/10.1016/j.micpro.2022.104593>
- Nathiya, N., Rajan, C. & Geetha, K.** (2025). A hybrid optimization and machine learning based energy-efficient clustering algorithm with self-diagnosis data fault detection and prediction for WSN-IoT application. *Peer-to-Peer Networking and Applications*, 18, Article 13. <https://doi.org/10.1007/s12083-024-01892-8>
- Prasad, K.H.V. & Periyasamy, S.** (2023). Secure–energy-efficient bio-inspired clustering and deep learning-based routing using blockchain for edge-assisted WSN environment. *IEEE Access*, 11, 145421–145440. <https://doi.org/10.1109/ACCESS.2023.3345218>
- Rathee, M., Kumar, S., Gandomi, A.H., Dilip, K., Balusamy, B. & Patan, R.** (2021). Ant colony optimization-based QoS-aware energy-balancing secure routing algorithm for Wireless Sensor Networks. *IEEE Transactions on Engineering Management*, 68(1), 170–182. <https://doi.org/10.1109/TEM.2019.2953889>
- Reddy, D.L., Puttamadappa, C. & Suresh, H.N.** (2021). Hybrid optimization algorithm for security-aware cluster head selection process to aid hierarchical routing in wireless sensor network. *IET Communications*, 15, 1561–1575. <https://doi.org/10.1049/cmu2.12169>
- Reddy, D.L., Puttamadappa, C. & Suresh, H.N.** (2021). Merged glowworm swarm with ant colony optimization for energy-efficient clustering and routing in Wireless Sensor Network. *Pervasive and Mobile Computing*, 71, 101338. <https://doi.org/10.1016/j.pmcj.2021.101338>
- Selvi, M., Kalaiarasi, G., Mana, S. C., Yogitha, R., & Padmavathy, R.** (2024). Energy and security aware hybrid optimal cluster-based routing in wireless sensor network. *Wireless Personal Communications*, 137(3), 1395–1422. <https://doi.org/10.1007/s11277-024-11288-9>
- Sharma, P., Sharma, M., Singh, R., Kumar, V., Agarwal, R. & Malik, P.K.** (2024). NHARSO-IWSN: A novel hybridized adaptive-network-based fuzzy inference system with reptile search optimization algorithm-based routing protocol for IoT-enabled Wireless Sensor Networks. *IEEE Transactions on Consumer Electronics*, 70(3), 6293–6302. <https://doi.org/10.1109/TCE.2024.3418845>
- Sheik Dawood, M., Sridevi, S., Akila, P. & Ramesh, J.** (2025). Hybrid Archimedes and arithmetic optimization algorithm for cluster head selection and multipath routing in Wireless Sensor Network. *International Journal of Communication Systems*, 38, e6100. <https://doi.org/10.1002/dac.6100>
- Srekanth, N., Babu, G.N., Kumar, G.R., G. M. & Shakir, A.M.** (2023). A cluster-based routing using energy-efficiency-based multi-objective optimization in Wireless Sensor Networks. In *2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics*, (pp. 1–6). IEEE. <https://doi.org/10.1109/AIKIE60097.2023.10389902>
- Thekiya, M.S. & Nikose, M.** (2022). Energy-efficient clustering routing protocol using novel admission allotment scheme (AAS) for intra-cluster communication in Wireless Sensor Network. *International Journal of Information Technology*, 14, 2815–2824. <https://doi.org/10.1007/s41870-022-01086-6>
- UmaRani, C., Ramalingam, S., Dhanasekaran, S., & Baskaran, K.** (2025). An hybrid machine learning and improved social spider optimization based clustering and routing protocol for wireless sensor network. *Wireless Networks*, 31(2), 1885–1910. <https://doi.org/10.1007/s11276-024-03861-8>
- Vanitha, M., Yamsani, N., Mohammed, I.H., Madhavan, S. & Al-Attabi, K.** (2023). Hybrid Salp Swarm and Particle Swarm Optimization based secure aware energy-efficient routing in Wireless Sensor Network. In *2023 International Conference on Integrated Intelligence and Communication Systems*, (pp. 1–5). IEEE. <https://doi.org/10.1109/ICIICS59993.2023.10421116>
- Wang, C., Liu, X., Hu, H., Han, Y., & Yao, M.** (2020). Energy-Efficient and Load-Balanced clustering routing protocol for wireless sensor networks using a chaotic genetic algorithm. *IEEE Access*, 8, 158082–158096. <https://doi.org/10.1109/access.2020.3020158>
- Wang, Z., Zeng, W., Yang, S., He, D. & Chan, S.** (2024). UCRTD: An unequally clustered routing protocol based on multihop threshold distance for Wireless Sensor Networks. *IEEE Internet of Things Journal*, 11(17), 29001–29019. <https://doi.org/10.1109/JIOT.2024.3406343>
- WSND.** (2025). Wireless Sensor Network Dataset. *Kaggle*. [https://www.kaggle.com/datasets/zia07/wireless-sensor-network-dataset?utm\\_source](https://www.kaggle.com/datasets/zia07/wireless-sensor-network-dataset?utm_source)
- Yassine, S., Najib, E.K. & Fatima, L.** (2019). Dynamic cluster head selection method for Wireless Sensor Network for agricultural application using fuzzy C-means clustering algorithm. In *2019 7th Mediterranean Congress of Telecommunications*, (pp. 1–9). IEEE. <https://doi.org/10.1109/CMT.2019.8931313>
- Zachariah, U.E. & Kuppusamy, L.** (2022). A hybrid approach to energy-efficient clustering and routing in Wireless Sensor Networks. *Evolutionary Intelligence*, 15, 593–605. <https://doi.org/10.1007/s12065-020-00535-0>