# A Novel Hybridization of ML Algorithms for Cluster Head Selection in WSN

Praveen Kumar R., M. P. Prabakaran, Durai Arumugam, and J. Selvakumar

***Abstract*—Generally, Wireless Sensor Networks (WSNs) are infrastructure-less networks with thousands of sensor nodes that sense or monitor the physical and environmental changes and forward the collected data to a central node. Besides, WSN has become the most efficient technology for handling Internet of Things (IoT) devices. Still, challenges such as node failures, high traffic among the nodes, link failures, etc., limit the performance of WSNs. To solve the challenges in WSN, this paper aims to develop a novel non-uniform clustering model, where the Cluster Heads (CHs) are selected based on the candidate CH selection strategy that transfers the data. Moreover, unbalanced energy utilization and data redundancy are eliminated via multi-hop communication. For attaining the non-uniform clustering model, the routing among the data packets is done by the efficiency of the hybridization of the Machin Learning (ML) algorithms viz Genetic Algorithm (GA) and Lion Algorithm (LA) with the consideration of energy, cost, time, network lifetime, and data accuracy. Finally, the performance of the proposed model is verified and validated through a comparative study with the existing models.**

***Index Terms*—Wireless Sensor Networks, Genetic Algorithm, Lion Algorithm, Cluster head Selection, Internet of Things**

## I. INTRODUCTION

Currently, IoT emerges as an incredible technology that enables anything that can be connected will be connected [1]. From a thing as tiny as a pill to a thing as large as an aircraft, it can be connected through IoT and collect and share data. With the innovation of digital intelligence, IoT connects various things associated with sensors and allows them to communicate without human interactions [2]. Besides, IoT virtually connects the world like a digital fabric that interlinks the digital and real universes. WSN generally consists of small sensor nodes managed and controlled through the base station (BS) [3]. Nevertheless, the battery capacity of the sensor nodes is limited by their size. Alternative power source deployment for every sensor node is practically infeasible since they are random and dynamic [4, 5]. In WSN, the entire network's performance relies on the lifetime of the sensor nodes, and the impairment resulting from dead nodes can drastically diminish the network's reliability. Accordingly, recent research is concentrating on the progress of innovative methodologies that minimize the nodes' vitality consumption [6-8].

Praveen Kumar R. is with the Dept. of ECE, Easwari Engineering College, Chennai, India. (E-mail: rpjcspraveen@gmail.com)

M. P. Prabakaran is with the Dept. of ECE, AKT Memorial College of Engineering and Technology, Kallakurichi, India.

Durai Arumugam is with the Dept. of IT, Easwari Engineering College, Chennai, India.

J. Selvakumar is with the Dept. of CSE, Sri Ramakrishna Engineering College, Coimbatore, India.

DOI: 10.36244/ICJ.2024.2.5

Also, WSN comprises many hubs, hubs management, and energy consumption strategies of every hub. Therefore, developing strategies to manage vitality accessibility is mandatory to ensure network lifetime [9, 10]. Only a few significant hubs are kept active through these topology control techniques to manage the availability and carry out the network capabilities [11]. The sensor nodes in the network are random and dynamic, as they may be either kept active or asleep. Hence, the hubs that possess maximal lasting energy are considered to be the significant ones [12, 13].

Rather than interconnecting every hub to sink a particular hub for transferring data, a pioneer node is nominated in every cluster concerning the node that possesses maximal lingering energy [14]. Other than the pioneer node/CH node, all other nodes are named as member nodes that transfer the data to the CH node [15]. At last, the gateway node establishes an appropriate connection among the lattices and transmits the data to the BS. Implementing clustering in WSN enables a typical communication bandwidth (BW), ensures an alleviated network, minimizes communication expenses, and eliminates data redundancy [16-19]. Previously, numerous types of research had been proposed to implement CH selection using ML and optimization concepts [19, 20]. Even though optimization strategies achieved better energy efficiency, network lifetime, and minimized delay, they need several parameters for delivering optimal solutions [21-24]. Moreover, choosing an appropriate algorithm to solve complex problems requires a better understanding and expertise in ML knowledge [25-27].

Although the rapid evolution of IoT and WSN offers unprecedented opportunities for network connectivity and data transmission, their potential is restricted by the limited lifespan of sensor nodes in the network. The ramifications of dead nodes causes reduced network performance, robustness and reliability. Hence, it is crucial to develop innovative technologies to reduce the energy consumption to ensure prolonged lifespan of sensor nodes. To achieve this objective, many existing techniques including Single-hop and Multi-hop CH Selection through GA [1], GA and Modified Particle Swarm Optimization (MPSO) [2], Fruitfly Optimization Algorithm (FFOA) and Glowworm Swarm Optimization (GSO) [3], New Individual Updating Strategies-based Hybrid Elephant Herding Optimization Algorithm [4], etc., are designed. These techniques aim to select the CH in the network to streamline communication, reduce energy consumption and eliminate data redundancy within the WSN. However, existing methodologies often rely on complex optimization techniques or machine learning algorithms, which demand large computational resources, parameter tuning and knowledge expertise to implement effectively. Moreover, they face certain challenges like reduction of alive nodes, increased cost function, less scalability, limited adaptability, etc. To overcome these issues, we proposed a promising solution by integrating evolutionary algorithms like GA and LA. The objective of our study is to enhance the longevity and energy efficiency of sensor nodes in WSNs through optimal CH selection. By selecting CHs, we aim to extend the lifespan of sensor nodes and improve overall energy efficiency across the network. The developed mechanism integrates the exploration and exploitation capabilities of GA and LA to effectively identify and select CHs in the WSN. This approach not only ensures improved energy efficiency but also offers reliable communication throughout the network. By combining these two meta-heuristic optimization algorithms into a single algorithm, this study provides a comprehensive solution that effectively addresses the challenges faced by existing CH selection methods in WSNs. The vital implementation descriptions of this paper are specified as follows.

- Initially, the significant CHs for the WSN are chosen through a hybridized adaptive EA model.
- The pioneer node in each lattice is identified and selected using the hybridization of GA and LA.
- To nominate the proper CH, the significant parameters are chosen regarding the constraints, viz., distance, energy, and delay.
- The performance of the proposed model is validated through a comparative investigation with other traditional models.

The rest of the paper is arranged in a fashion as given below. Section II reviews the literature concerning various CH selection protocols using different approaches. Section III deliberates the proposed architecture and its objectives. Further, Section IV explains the conventional and proposed optimization concepts used. The simulation results and the achievements are discussed in Section V.

## II. Literature Study

In 2021, Praveen Kumar R *et al.* [1] developed a CH selection model for WSN using single-hop and multi-hop CH selection through GA (S/MHCH-GA). This work implemented a novel CH selection to attain significant performance by exploiting both single- and multi-hop criteria. Moreover, the S/MHCH-GA selected CH based on the energy constraint by estimating the initial and utilized energy. The empirical results proved that the S/MHCH-GA model attained an energy-efficient WSN model and outperformed the existing methods. However, this approach is limited adaptability to dynamic environments.

Praveen Kumar R et al. [2] have introduced a CH selection model for WSN using hybridization of optimization algorithms, namely GA and Modified Particle Swarm Optimization (MPSO). Herein, the main aim of this work was to attain an energy-efficient and network lifetime-enhanced WSN model by optimizing the parameters of energy utilization, delay, throughput, network lifespan, and energy efficacy. Finally, the performance of the proposed method was validated with a comparison of the conventional optimization algorithms. However, this framework requires fine-tuning of algorithm parameters leading to high time-consuming and resource-intensive.

In 2019, Kale Navnath Dattatraya and K. Raghava Rao [3] presented a fitness-based Glowworm swarm with the Fruitfly approach (FGF) for CH selection in WSN. The implementation used a hybridized strategy of the Fruitfly Optimization Algorithm (FFOA) and Glowworm Swarm Optimization (GSO). The significant parameters considered for optimization were network lifespan, energy, delay, and cost. The simulation investigation shows the performance of the proposed FGF method by outperforming the traditional optimization methods. However, this hybrid methodology faces issues like high complexity, and it is not scalable to handle large WSN.

In 2020, Sim Sze Yin and Yoni Danieli [4] developed a New Individual Updating Strategies-based Hybrid Elephant Herding Optimization Algorithm (NIUS-HEHOA) model for CH selection in WSN. It began with clustering, energy utilization estimation, and enhancing network lifespan. The obtained results showed the performance of the NIUS-HEHOA, which accomplished better throughput, alive nodes, and energy than other conventional methods. Although this framework achieved improved results than the conventional models, it lacks adaptability to the changing dynamic characteristics of WSN, leading to performance reduction in real-time application.

In 2018, A. Rajagopal *et al.* [5] addressed a CH selection model for WSN using the hybridization of Bacterial foraging Optimization (BFO) and bee swarm Optimization (BSO). However, the main objective was to minimize the packet delivery ratio and maximize energy efficiency. The performance of the hybrid BFO-BSO approach was validated in terms of delay, energy, and lattice count. However, the CH selection using this hybrid approach is time-consuming and resource-constraint, making it less applicable for real-world WSN scenarios.

In 2021, Umashankar ML *et al.* [6] adopted a hybrid Simulated Annealing (SA) approach to nominate efficient pioneer nodes for WSN. The aim was to attain energy efficiency by considering battery capabilities like size, rechargeable, and replaceable properties. The investigation analysis revealed better network lifespan and robustness

TABLE I
LITERATURE SURVEY

| Authors | Methodology | Benefits | Drawbacks |
|---|---|---|---|
| Praveen Kumar R et al. [1] | Single-hop and Multi-hop CH Selection through GA | Energy-efficient CH selection, suitable for resource-constrained environment | Limited adaptability and scalability to dynamic environments |
| Praveen Kumar R et al. [2] | GA and Modified Particle Swarm Optimization (MPSO) | Optimizes energy utilization, throughput, network lifespan, and energy efficacy | Requires fine-tuning and time-consuming |
| Kale Navnath Dattatraya and K. Raghava Rao [3] | Fruitfly Optimization Algorithm (FFOA) and Glowworm Swarm Optimization (GSO) | Increases network lifespan, and minimizes delay and cost function | Complex and cannot handle large networks |
| Sim Sze Yin and Yoni Danieli [4] | New Individual Updating Strategies-based Hybrid Elephant Herding Optimization Algorithm | Energy optimization, improves network lifespan, increased throughput and alive nodes | Cannot adapt to the dynamic changes in WSN, making it inefficient for real-time application |
| A. Rajagopal et al. [5] | Bacterial foraging Optimization (BFO) and bee swarm Optimization (BSO) | Improves packet delivery ratio and maximizes energy efficiency | Time-consuming and resource-intensive |
| Umashankar ML et al. [6] | Simulated Annealing (SA) algorithm | Achieved improved energy efficiency, and better network lifespan than other conventional models | Increased computational overhead and limited reliability |

through customized battery enhancements. However, this method faces challenges like high computational overhead, and limited reliability. Table 1 presents the literature survey.

In the related work section, we reviewed several existing algorithms for CH selection in WSN. Each technique offers certain advantages and disadvantages. The major drawbacks of the existing approaches include limited scalability, less adaptability, high computational overhead, demand of fine-tuning, large time consumption, etc. Moreover, some methods are resource-intensive, making it less applicable for real-world CH scenarios, since the WSN are resource-constrained environments. In addition, the existing works face difficulty in integrating the different optimization algorithms into a single approach, leading to increased system complexity. To resolve these issues, we proposed a novel CH selection protocol using the combination of Genetic Algorithm and Lion Algorithm. The developed methodology aims to improve the performance of WSN by enhancing the CH selection efficiency. This algorithm considers the constraints like distance, energy, delay, etc., for selecting the CH. Unlike the existing works discussed in the literature survey, which predominantly focus on single optimization algorithms or their hybrids, the designed approach strategically combines GA and LA to exploit their complementary strengths, enhancing the efficiency and effectiveness of CH selection. The improved exploration and exploitation phase of the developed algorithm enables to identify the optimal CH candidate in the solution space. In addition, the developed algorithm prioritizes CH selection nodes with high energy, offering robust and more reliable communication within the WS network. Moreover, the iterative CH selection process ensures prolonged network lifespan and increases the energy utilization, improving the WSN performances. Thus, the proposed algorithm addresses the challenges faced by the conventional models.

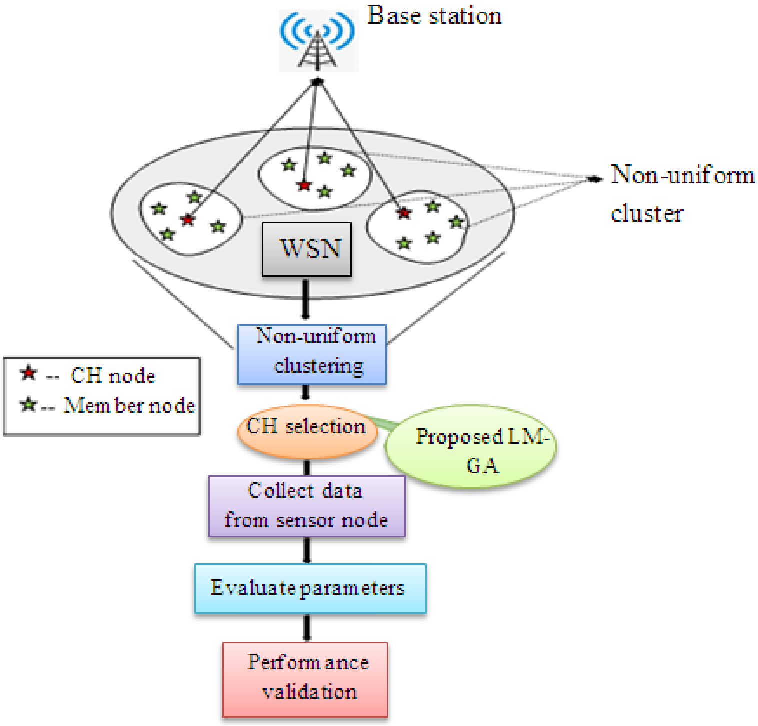



Fig. 1. Block Diagram of Proposed CH Selection Model for WSN

## III. A Novel CH Selection Model for WSN

### A. *Proposed Architecture*

This section discusses the novel CH selection protocol established to accomplish the desired WSN performance. Fig. 1 depicts the model of the proposed CH selection protocol. Initially, the significant CHs for the WSN are chosen through a hybridized adaptive EA model. The hybridized adaptive Evolutionary Algorithm (EA) model uses the combination of Genetic Algorithm (GA) and Lion Algorithm (LA) for optimal selection of CHs in WSN[32, 33]. The GA is an evolutionary optimization algorithm developed based on the concept of genetics. This approach iteratively evolves a population of candidate solutions to estimate the optimal solution by following the steps like selection, mutation, crossover, etc. Consequently, the LA model aims to refine solutions iteratively; thereby improving the models CH process. The pioneer node in each lattice is identified and selected using the hybridization of GA and LA. The significant parameters about the constraints, viz distance, energy, and delay, are chosen to nominate the proper CH. Here, a non-uniform clustering approach is employed to cluster the network, which also assists in minimizing network cost and battery costs. Thereby, a cost-effective system is achieved. Now, non-uniform clustered lattices are created based on the CH nodes' energy efficiency, and the cluster's other nodes are collectively termed member nodes.

The CH nodes are selected through the proposed, and the election process iteratively takes place. According to the fitness of the proposed model, the CH nodes are evaluated based on the residual energy for each iteration. Each sensor node's energy drains through data transmission over the network by choosing an optimal CH that is reliable as far as efficient WSN is possible. The proposed model effectively carries out this process. The energy $E$ estimation for every data transmission is stated in Eq. (1), where $E(Q)$ specifies the energy among the CH and member nodes, and the estimation of $E(Q)$ is given in Eq. (2). Besides, $U(i)$ is the energy between CH and BS in Eq. (3). Further, $E(P)$ is the vitality among two member nodes and the calculation is explained in Eq. (4), in which $K(N_j)$ is the $j^{th}$ member node energy, $K(CH_i)$ is the $i^{th}$ CH energy, $n$ speaks out the CH count, and $l$ points to the total number of nodes in the network.

$$E = \frac{E(Q)}{E(P)} \tag{1}$$

$$E(Q) = \sum_{i=1}^{n} U(i) \tag{2}$$

$$U(i) = \sum_{\substack{j=1 \\ j\in i}}^{l} (1 - K(N_j) \times K(CH_i) : 1 \leq i < n \tag{3}$$

$$E(P) = m \times max_{j=1}^{l} K(N_j) \times max_{i=1}^{n} K(CH_i) \tag{4}$$

The threshold distance after every iteration is calculated according to Eq. (5), which is the distance derivation among the nodes in the network. Here, $BS$ is the base station, $D(Q)$ is the distance between the CH and member nodes as defined in Eq. (6), and $D(P)$ refers to the distance between two adjacent member nodes as presented in Eq. (7).

$$D = \frac{D(Q)}{D(P)} \tag{5}$$

$$D(Q) = \sum_{j=1}^{l} \sum_{i=1}^{n} \left\| N_j - CH_i \right\| + \left\| CH_i - BS \right\| \tag{6}$$

$$D(P) = \sum_{j=1}^{l} \sum_{i=1}^{n} \left\| N_j - N_i \right\| \tag{7}$$

Moreover, the network lifespan is directly proportional to the amount of data transferred. Reducing data transmission delay can improve the network lifespan. Eq. (8) shows the delay due to data transfer among the nodes, where $max_{i=1}^{n}(CH_i)$ is the total number of CHs.

$$D_t = \frac{max_{i=1}^{n}(CH_i)}{l} \tag{8}$$

These are the significant parameters identified to influence the WSN performance.

## IV. TRADITIONAL AND PROPOSED GA AND LA MODELS

### A. Traditional GA

Like all other optimization concepts, GA [29] begins with population initialization, fitness estimation, and fitness. The termination of GA is carried out by convergence like other optimizations as well. Nevertheless, the operations in between the GA varies with other optimization concepts. Generally, GA comprises $S$ chromosomes with population $R$, and fitness is estimated for each chromosome. Following fitness, mate selection takes place to perform crossover and mutation. These are the main three steps in GA. The mathematical model of GA is expressed in Eq. (10), where $P(a_i)$ is the probability of selected individual $S$.

$$P(a_i) = \frac{f(a_i)}{\sum_{j=1}^{n} f(a_i)} \tag{10}$$

In many cases, the $P(a_i)$ chosen as unmodified from two-parent chromosomes $N$ produces no new solutions. Moreover, the child's genes modify arbitrarily, and mutation is normally slow in GA. The steps in GA optimization are described below,

***Population initialization:*** The GA optimization begins with the initialization of the population of potential solutions with fixed population size. The population size defines the number of solutions in each generation of the GA optimization.

***Selection:*** Determine the fitness function for each individual in the population. The fitness function is evaluated based on objective function (optimization problem). After fitness evaluation, the individuals or solutions with higher fitness were selected for reproduction.

***Crossover:*** In this phase, the selection solutions are combined to create new offspring through crossover operation. This phase involves exchanging the genetic materials between the selected individuals to create new solutions.

***Mutation:*** After crossover step, the next process is mutation in which random changes are done to the created new solutions to prevent premature convergence to suboptimal solutions.

***Termination:*** The final step is termination. The GA algorithm iterates through generations until a termination criterion is met. The termination conditions include maximum number of iterations, maximum convergence, etc.

### B. Traditional LA

LA [28] is developed based on the biological life behavior of lions. Here, the male lion is $R_{male}$, the female lion is $R_{female}$, and the nomad lion is $R_{nomad}$. There are two main processes, namely crossover and mutation, and one auxiliary process, such as gender clustering, is concerned with mating. Hither, $R_{male}$, and $R_{female}$ produce up to four cubs after computing crossover. Furthermore, cubs are the solutions attained by both $\square_{male}$s and $\square_{female}s$. Every cub is delivered through a different crossover mask $C_m$. In other words, $m^{th}$ mask $C_m$ is utilized for producing $R^{cubs}(m)$. Further, these cubs $R^{cubs}$ are subjected to mutation, and produce another four cubs namely $R^{new}$. Subsequently, all these cubs are placed in the cubs' pool, and the gender clustering process takes place to decide $R^{m\_cubs}$ and $R^{f\_cubs}$. The steps involved in LA are described below,

***Initialization of Population:*** The LA optimization commences with the initialization of lions with fixed population size.

***Position Initialization:*** *Assign and initialize the position of each lion in the population. The position indicates the potential candidate solution for solving the optimization problem.*

***Fitness Evaluation:*** After initialization, determine the fitness of each lion in the population.

***Update Leader:*** Select the lion with higher fitness value, and it is considered as leader. It guides the pride in the search process.

***Pride Update:*** Then, based on the leader position, the positions of the pride are updated. This phase involves updating the lion's position to explore the search space around the leader's position.

***Boundary Handling:*** In this step, the updated positions of the lions are verified whether it lies within the search space boundary. In case any lion deviated from the boundaries, its position was adjusted.

***Hunting Behavior:*** Further, simulate the lion's hunting mechanism. Here, the lions explore the search space to find the potential prey.

***Prey Capture and Sharing:*** On finding the prey, the lion shares the information with other lions in the population. The process aids in exchanging the information between the pride and leader, guiding the lions towards optimal solutions.

***Termination Criteria:*** Check for termination conditions like maximum number of iterations or maximum convergence. If the termination is not met, repeat the steps 3 to 9.

## V. SIMULATION RESULTS

### A. Simulation Setup

The proposed CH selection for WSN using the hybrid approach was implemented in MATLAB 2018a on Intel core® core i3 processor 7020U@2.3 GHz, 8 GB RAM, 64-bit operating system. The efficiency and novelty of the implemented model were recorded via the simulation results. Hither, the network nodes spread over $100m \times 100m$ and amidst a BS. The initial vitality $V_{init} = 0.5$, free space vitality $V_{free\_s} = 10pj/bit/m^2$, power amplifier vitality $V_{power} = 0.0013pj/bit/m^2$, transmitter amplifier vitality $V_{trans} = 50nj/bit/m^2$, and data aggregation factor $V_{da} = 5nj/bit/signal$, and the total rounds count is 2000. The evaluation was implemented through various performance parameters such as energy, distance, delay, network lifetime, cost, and network accuracy. Furthermore, the performance of the proposed model is compared over various conventional models such as LA [28], GA [29], Grey Wolf Optimization (GWO) algorithm [30], and Whale Optimization Algorithm (WOA) [31].

### B. Algorithmic Analysis

In this subsection, the performance of the proposed CH selection for WSN using the hybrid proposed model is discussed. The statistical analysis concerning mean, median, and standard deviation (SD) for proposed and traditional approaches concerning the number of alive nodes and normalized energy are given in Table II. The analysis clearly shows that the proposed model attained better energy efficiency. Figure 2 depicts the evaluation of the number of alive nodes over increasing iterations. This metric indicates the number of sensor nodes that are functioning within the network over the iterations. It enables to determine the network's resilience and robustness against node failures.

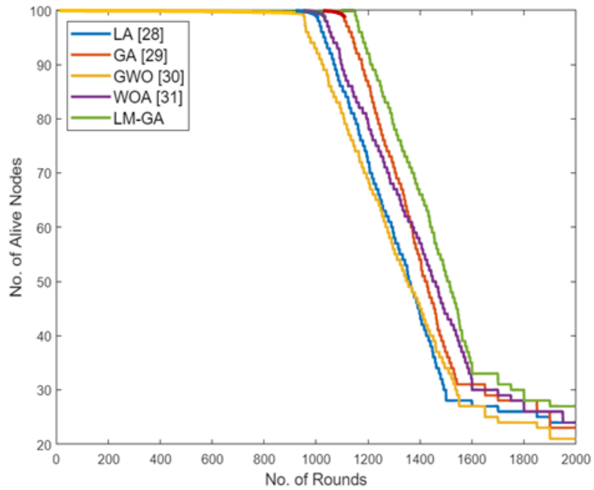



Figure 2: Evaluation of number of alive nodes

Here, the number of alive nodes are assessed over increasing iterations from 0 to 2000. The number of alive nodes achieved by the proposed method is compared and evaluated with the existing techniques such as GA, LA, GWO, and WOA. From the analysis, it is observed that the number of alive nodes in the existing techniques as well as the proposed model is constant that is 100 from 0 to 1000 iterations. After 1000 iterations, the number of alive nodes starts decreasing. At 2000 iterations, the number of alive nodes in conventional models like GA, LA, GWO, and WOA is 21, 25, 23, and 24, while the number of alive nodes in the proposed algorithm is 30. This analysis validates that the designed approach has the highest number of alive nodes than the existing models. This efficiency of the designed approach offers several advantages like improved network coverage, enhanced data delivery, and prolonged network lifetime.

TABLE II
THE STATISTICAL ANALYSIS CONCERNING MEAN, MEDIAN, AND SD FOR PROPOSED AND TRADITIONAL APPROACHES

| ***Approaches*** | | GA [29] | LA [28] | GWO [30] | WOA [31] | Proposed model |
|---|---|---|---|---|---|---|
| Mean | Alive Nodes | 62.7 | 62.4 | 62.3 | 62.5 | 62.9 |
| | Normalized Energy | 0.20154 | 0.20187 | 0.20114 | 0.20199 | 0.20214 |
| Median | Alive Nodes | 75 | 77 | 77 | 73 | 79 |
| | Normalized Energy | 0.11547 | 0.11487 | 0.11211 | 0.11114 | 0.11747 |
| SD | Alive Nodes | 41.214 | 41.747 | 41.457 | 41.545 | 41.987 |
| | Normalized Energy | 0.20878 | 0.20854 | 0.20877 | 0.20477 | 0.20114 |

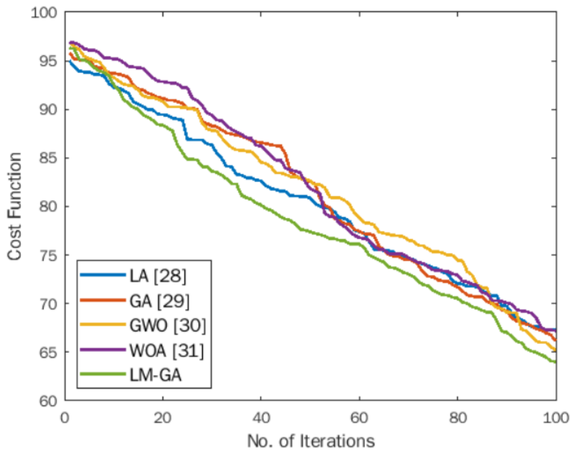



Figure 3: Evaluation of cost function

Figure 3 presents the comparison of cost function. The cost function is the metric, which measures the suitability of the sensor nodes for assuming the role of cluster heads. The cost function enables to assess multiple parameters like energy levels, communication range, energy consumption, communication overhead, and computational capabilities. Here, the cost function is compared with the existing techniques such as GA, LA, GWO, and WOA. The cost function is assessed over varying number of iterations from 0 to 100. From the evaluation, it is observed that the proposed algorithm obtained a minimum cost function compared to other techniques. At 100 iterations, the conventional models and proposed approach attained cost function of 68, 66, 70, 66, and 63, respectively. The minimal cost function indicates that the proposed algorithm contributes to reduced energy consumption, enhanced network coverage, efficient data aggregation, improved network performances, etc.

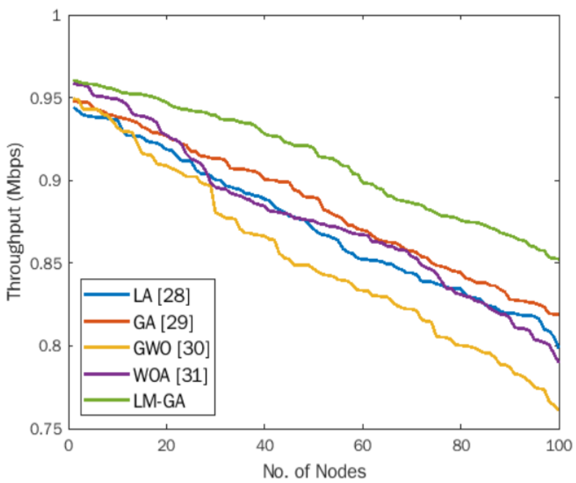



Figure 4: Throughput validation

Figure 4 depicts the validation of the throughput. The throughput metric measures the rate at which data packets are successfully transmitted from source to destination within the network. We measured the throughput in terms of megabits per second (Mbps). The higher throughput indicates improved network efficiency in transmitting data. Here, we evaluated the network throughput by varying the number of nodes from 0 to 100. The determined throughput of the developed algorithm is compared with the conventional models such as GA, LA, GWO, and WOA. From the analysis, it is observed that when the number of nodes increases, the network throughput decreases. When the number of nodes is 100, the network throughput of the above-stated existing techniques and proposed approach is 0.76, 0.78, 0.82, 0.8, and 0.86, respectively. This manifests that the developed algorithm improved the network throughput than the existing techniques. In addition, the comparative assessment highlights that the developed algorithm acts as a promising solution for improving network efficiency, enhancing data delivery, and facilitating reliable communication in wireless sensor networks.

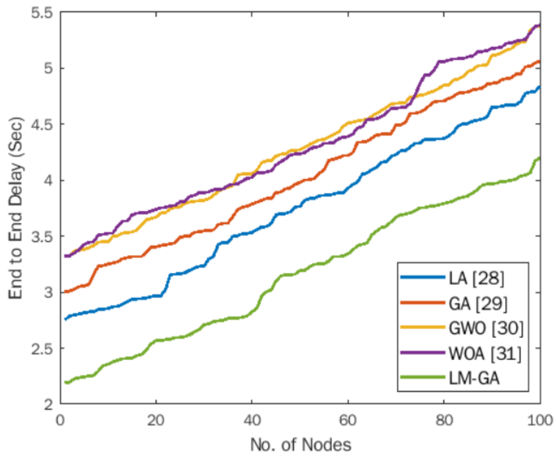



Figure 5: End-to-End delay

Figure 5 depicts the end-to-end delay. The end-to-end delay indicates the time taken for a data packet to travel from source to destination node. This metric includes all delays including transmission delay, propagation delay, queuing delay, and processing delay. To validate that the presented algorithm obtained minimum end-to-end delay, it is compared with existing techniques such as GA, LA, GWO, and WOA. Here, the end-to-end delay was assessed over increasing node count in the network. This evaluation states that the end-to-end delay increases on increasing the node count. When node count is 100, the above-mentioned conventional models and the proposed algorithm obtained end-to-end delay of 5.1s, 4.9s, 4.6s, 5.3s, and 3.8s, respectively. This lower delay implies that the developed algorithm delivers the data quickly with minimal latency in the network.

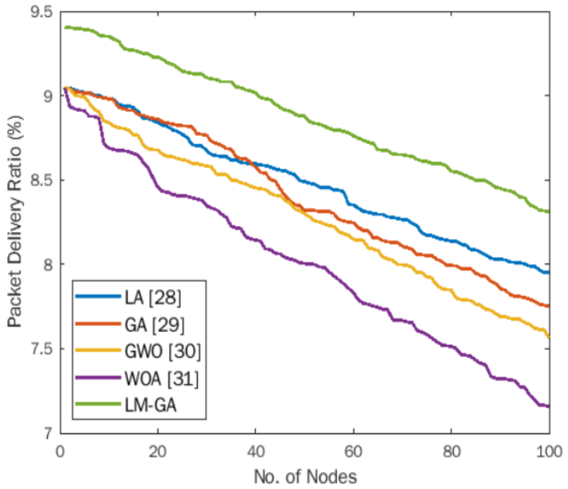



Figure 6: PDR validation

Figure 6 presents the validation of the packet delivery ratio (PDR) with the existing models. the PDR defines the proportion of the number of data packets successfully received at the destination to the total number of data packets sent by the source. It represents the effectiveness of data delivery in the network. The traditional models such as GA, LA, GWO, and WOA earned PDR of 7.2%, 8.2%, 7.6%, and 7.9%, while the developed methodology attained PDR of received at the destination to the total number of data packets sent by the source. It represents the effectiveness of data delivery in the network. The traditional models such as GA, LA, GWO, and WOA earned PDR of 7.2%, 8.2%, 7.6%, and 7.9%, while the developed methodology attained PDR of 8.5%. This illustrates that the developed algorithm offers better reliability and network performances than the existing techniques. Also, it indicates the data delivery efficiency of the network.

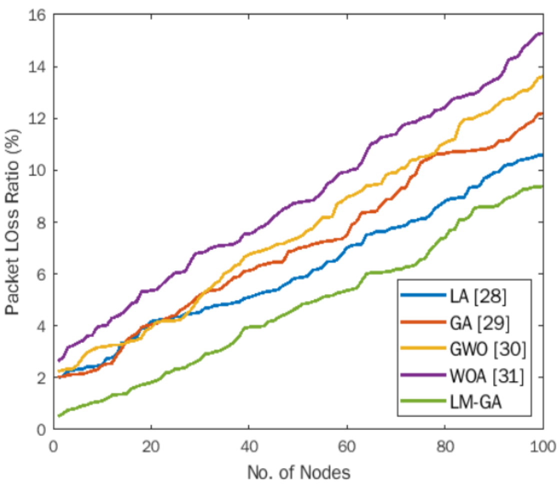



Figure 7: PLR validation

Consequently, we evaluated the packet loss ratio (PLR), which was depicted graphically in Figure 7. The packet loss ratio indicates the proportion of the number of lost packets during transmission to the total number of packets sent. It measures the reliability of the system to deliver data in WSN. The PLR is assessed by increasing the number of nodes in the network to determine the scalability of the proposed algorithm. The existing models including GA, LA, GWO, and WOA achieved PLR of 15%, 13%, 12%, and 10%, while the developed methodology achieved PLR of 9% for 100 nodes. The reduction in proposed model's PLR suggests that it is reliable in transmitting the data packets to the destination. Also, the lower PLR indicates that only few packets are lost during transmission, indicating the better network performance.

The energy consumption measures the amount of energy consumed by the sensor nodes to perform tasks like communication, data processing, sensing, etc., within the WSN. The reduction of energy consumption increases the network lifetime and improves the overall network efficiency. Figure 8 depicts the comparison of energy consumption. Here, the energy consumption metric was evaluated by increasing the number of nodes from 0 to 100. When node count is 100, the energy consumption obtained by the existing techniques like GA, LA, GWO, and WOA is 0.9mJ, 0.85mJ, 0.75mJ, and 0.8mJ, respectively, while the proposed framework consumed minimum energy of 0.55mJ. This minimum energy consumption of the developed framework highlights its applicability in resource-constrained WSN than the existing models. Also, its reduction of energy consumption aids in minimizing the cost function.

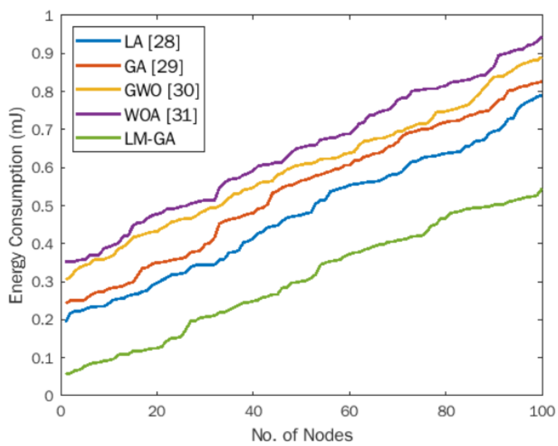



Figure 8: Energy consumption comparison

From this intensive evaluation of proposed model's performances with the conventional models like GA, LA, GWO, and WOA, it is evident that the developed model achieved improved outcomes in terms of PDR, throughput, achieved improved outcomes in terms of PDR, throughput, and number of alive nodes. On the other hand, the metrics such as energy consumption, packet loss ratio, end-to-end delay, and cost function attained by the proposed technique are less compared to other models. This comprehensive performance evaluation suggests that the proposed method's effectiveness and reliability in providing improved Quality of Service in the WS network. Thus, the proposed model achieved considerable performance over the conventional approach and proved its efficiency.

## VI. Conclusion

This paper has established a novel ML algorithm that hybridizes optimization concepts with an EA model for energy-efficient CH selection for WSNs. Initially, CHs for every lattice were selected based on the candidate CH selection strategy, which transfers the data. Moreover, unbalanced energy utilization and data redundancy were eliminated via multi-hop communication. For attaining the non-uniform clustering model, the routing among the data packets was done by the efficiency of the hybridized model considering energy, cost, time, network lifetime, and data accuracy. Finally, the performance of the proposed model was verified and validated through a comparative study with the existing models such as WOA, GWO, LA, and GA. The throughput achievement of the proposed model obtained 7.6%, 8.69%, 8.71%, and 13.04% better than LA, GA, GWO, and WOA, respectively. Also, the packet loss ratio attained 23.33%, 29.14%, 37.455, and 39.74% improved than LA, GA, GWO, and WOA, respectively. Thereby, the proposed model performed well and outperformed the conventional approaches. Moreover, the developed framework involves making strategic routing decisions, it ensures tradeoff between the high network lifespan and minimum delay. Also, this framework achieved an optimal balance between the data transmission accuracy and energy efficiency, making it an effective protocol for real-world CH selection scenarios.

Although the proposed work achieved better results, it faces certain limitations. Firstly, the developed algorithm is sensitive to network conditions like node mobility, charging environmental factors, topology variations, etc. These fluctuations decrease the stability and reliability of the proposed framework. Secondly, although the developed approach offered higher scalability than the existing works, it still faces problems in enhancing the scalability to meet the real-time demand in WSN. To overcome these issues, the future study should concentrate on developing adaptive models with continuous learning to resolve these issues.

## References

[1] Praveen Kumar R, Jennifer S. Raj, and Smys S, "Analysis of dynamic topology wireless sensor networks for the Internet of Things (IoT), " International Journal of Innovations in Engineering and Technology (IJIET), Sept 14 2021.

[2] Praveen Kumar R, Jennifer S. Raj, and Smys S, "Performance Analysis of Hybrid Optimization algorithm for Cluster Head Selection in Wireless Sensor Networks, " in press.

[3] Kale Navnath Dattatraya and K. Raghava Rao, "Hybrid-based cluster head selection for maximizing network lifetime and energy efficiency in WSN, " Journal of King Saud University - Computer and Information Sciences, Available online Apr 4 2019.

[4] Sim Sze Yin, and Yoni Danieli, "A Hybrid Optimization Algorithm on Cluster Head Selection to Extend Network Lifetime in WSN," Journal of Computing in Engineering, pp. 7–18, 2020.

[5] A. Rajagopal, S. Somasundaram, B. Sowmya, Performance Analysis for Efficient Cluster Head Selection in Wireless Sensor Network Using RBFO and Hybrid BFO-BSO, International Journal of Wireless Communications and Mobile Computing. Vol. 6, No. 1, 2018, pp. 1–9.

[6] Umashankar ML, Anitha TN, and Mallikarjunaswamy S. (2021) An efficient hybrid model for cluster head selection to optimize wireless sensor network using simulated annealing algorithm. Indian Journal of Science and Technology. Vol. 14, No. 3, pp. 270–288.

[7] Rajagopal, A., S. Somasundaram and B. Sowmya. "Performance Analysis for Efficient Cluster Head Selection in Wireless Sensor Network Using RBFO and Hybrid BFO-BSO." International Journal of Wireless Communications and Mobile Computing, Vol. 6, No.1, March 2018, pp. 1–9.

[8] N. Lavanya, and T. Shankar, "Hybrid based Energy Efficient Cluster Head Selection using Camel Series Elephant Herding Optimization Algorithm in WSN," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 5, 2020.

[9] J Sathya Priya, Dr. Wael Mohammad Alenazy, and A. R Sathyabama, "Energy Efficient Cluster based Routing Protocol for WSN based on Hybrid BSO-TLBO Optimization Model," Research Square, June

[10] Er. Pankaj Bajaj, 2017, Efficient Cluster Head Selection in Wireless Sensor Networks using Bacteria Foraging Optimization, International Journal Of Engineering Research & Technology (IJERT) NCIETM, Vol. 5, No. 11, 2017.

[11] Rakesh Kumar Yadav, and Rajendra Prasad Mahapatra, "Energy-Aware Optimal Cluster Head Selection Using Hybrid Algorithm for Clustering Routing in Wireless Sensor Networks," International Journal of Intelligent Engineering and Systems, Available Online, Jan 11, 2019.

[12] T. M. Behera, S. K. Mohapatra, U. C. Samal, M. S. Khan, M. Daneshmand and A. H. Gandomi, "Residual Energy-Based Cluster-Head Selection in WSNs for IoT Application," IEEE Internet of Things Journal, vol. 6, no. 3, pp. 5132–5139, June 2019.

[13] A. Ali Y. Ming T. Si S. Iram and S. Chakraborty "Enhancement of RWSN lifetime via firework clustering algorithm validated by ANN" Information vol. 9 no. 3 p. 60 2018.

[14] J. Duan D. Gao D. Yang C. H. Foh and H.-H. Chen "An energy- aware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications" IEEE Internet Things J. vol. 1, no. 1, pp. 58–69, Feb. 2014.

[15] P. C. S. Rao P. K. Jana and H. Banka "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks" Wireless Netw. vol. 23, no. 7, pp. 2005–2020, 2017.

[16] H. Farman et al. "Multi-criteria based zone head selection in Internet of Things based wireless sensor networks" Future Gener. Comput. Syst. vol. 87, pp. 364–371, Oct. 2018.

[17] S. Kallam R. B. Madda C.-Y. Chen R. Patan and D. Cheelu "Low energy-aware communication process in IoT using the green computing approach" IET Netw. vol. 7, no. 4, pp. 258–264, Aug. 2017.

[18] H. Farman H. Javed J. Ahmad B. Jan and M. Zeeshan "Grid-based hybrid network deployment approach for energy-efficient wireless sensor networks" J. Sensors vol. 2016, 2016.

[19] S. Umbreen, D. Shehzad, N. Shafi, B. Khan and U. Habib, "An Energy-Efficient Mobility-Based Cluster Head Selection for Lifetime Enhancement of Wireless Sensor Networks," IEEE Access, vol. 8, pp. 207 779–207 793, 2020.

[20] K. A. Darabkh and L. Al-Jdayeh "A new fixed clustering based algorithm for wireless sensor networks" Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. pp. 71–76, Jun. 2018.

[21] K. Guleria and A. K. Verma "Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks" Wireless Netw. vol. 25 no. 3 pp. 1159–1183 Apr. 2019.

[22] H. El Alami and A. Najid "ECH: An enhanced clustering hierarchy approach to maximize the lifetime of wireless sensor networks" IEEE Access vol. 7, pp. 107 142–107 153, 2019.

[23] Y. H. Robinson E. G. Julie R. Kumar and L. H. Son "Probability-based cluster head selection and fuzzy multipath routing for prolonging lifetime of wireless sensor networks" Peer Netw. Appl. vol. 12 no. 5, pp. 1061–1075, May 2019.

[24] H. Mostafaei "Energy-efficient algorithm for reliable routing of wireless sensor networks" IEEE Trans. Ind. Electron. vol. 66 no. 7, pp. 5567–5575, Jul. 2019.

[25] M. Elhoseny and A. E. Hassanien "Hierarchical and clustering WSN models: Their requirements for complex applications" in Dynamic Wireless Sensor Networks Cham Switzerland: Springer vol. 165 2019.

[26] S. Yahiaoui M. Omar A. Bouabdallah E. Natalizio and Y. Challal" An energy-efficient and QoS aware routing protocol for wireless sensor and actuator networks" AEU—Int. J. Electron. Commun. vol. 83, pp. 193–203, Jan. 2018.

[27] Z. Zhao K. Xu G. Hui and L. Hu "An energy-efficient clustering routing protocol for wireless sensor networks based on AGNES with balanced energy consumption optimization" Sensors vol. 18, no. 11., p. 3938, Nov. 2018.

[28] Rajakumar Boothalingam, "Optimization using lion algorithm: a biological inspiration from lion's social behavior," Evolutionary Intelligence, Vol. 11, pp. 31–52, 2018.

[29] Yi Ding, and Xian Fu, "Kernel-Based Fuzzy C-Means Clustering Algorithm Based on Genetic Algorithm," Neurocomputing, Vol. 188, May 5 2016, pp. 233–238.

[30] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewisa, "Grey Wolf Optimizer," Advances in Engineering Software, Vol. 69, pp. 46–61, March 2014.

[31] Seyedali Mirjalili, and Andrew Lewis, "The Whale Optimization Algorithm", Advances in Engineering Software, Vol. 95, pp. 51–67, May 2016.

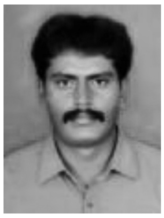
[32] Galán-Jiménez, J. (2017, December). Minimization of energy consumption in IP/SDN hybrid networks using genetic algorithms. In 2017 Sustainable Internet and ICT for Sustainability (SustainIT) (pp. 1–5). IEEE.

[33] Phillips, C., Gazo-Cervero, A., & Chen, X. (2011, October). Pro-active energy management for wide area networks. In IET International Conference on Communication Technology and Application (ICCTA 2011) (pp. 317–322). IET.

**R. Praveen Kumar**, obtained his B.E., from Seethai Ammal Engineering College and M.E., from Sona College of Technology, Anna University, He received his PhD degree from Anna University, India, Chennai. for his research work on Internet of Things.

He is currently a full Associate Professor at Easwari Engineering College, India, Chennai. He is the author and coauthor of more than 25 papers published in prestigious journals and conference proceedings.

R. Praveen Kumar is a member of the IEEE, Fellow member in IETE, Life member in IE. His research interests include Wireless Sensor Networks, Internet of Things, Wireless Networks. To his credit, he has one granted Australian patent and One Indian Patent (Published).



**M. P. Prabakaran** is currently working as an Associate Professor in the Department of Electronics and Communication Engineering at A.K.T Memorial College of Engineering in Kallakurichi, Tamil Nadu, and India. He received his undergraduate degree in Electronics and Communication Engineering in 2006, as well as his Master of Engineering degree in 2008 from Anna University, Tamil Nadu. He was awarded a gold medal for achieving first rank in his master's program. In 2019, he obtained his PhD in Optical Wireless Communication from Anna University, Tamil Nadu. Dr. M.P. Prabakaran has published several papers in reputable journals and contributed chapters to books. Additionally, he has presented various academic and research papers at national and international conferences. His current research activities focus on optical wireless communication, signal processing, and wireless sensor networks.

**Durai Arumugam** S S L received the B.Tech. Degree in Information Technology and the M.E. degree in Software Engineering from Anna University, Chennai, India. Currently pursuing Ph.D. in the area of Deep Learning in Anna University, Chennai. He has more than 9 years of teaching and research experience. He is currently working as an Assistant Professor at Easwari Engineering College, Chennai in the Department of Information Technology. Had 10 publications in reputed journals and International Conferences. His current research interests include machine learning and data analytics. He has organized a few conferences, including one virtual conference.

**J. Selvakumar** completed his Ph.D from Anna University, Chennai in March 2013, Post Graduation in M.E (CSE) from Kongu Engineering College in 2002 and Under graduation in B.E (CSE) from Madras University in 2001. He has more than 21 Years of Teaching and Research Experience. He is presently working as a Professor in Computer Science & Engineering Department in Sri Ramakrishna Engineering College, Coimbatore. He has around 82 National/International Conference and Journal Publications. His research area are Software Engineering, Agile, Dev Ops, Full Stack Development, Green Computing, Orange Computing, Carbon Footprint and Ground Water Recharge Technology. He is IEEE Senior Member (Membership No : 95408388). He has received Research grants from AICTE, Anna University and IEEE. He has also published a Patent.