# Particle Swarm Optimization Based Energy Efficient Clustering and Sink Mobility in Heterogeneous Wireless Sensor Network

**Biswa Mohan Sahoo[a]  Tarachand Amgoth[b]  Hari Mohan Pandey[c]**

Amity University, Greater Noida, India
Indian Institute of Technology (Indian School of Mines), Dhanbad
Edge Hill University, Ormskirk, Lancashire, UK
bmsahoo@gn.amity.edu, tarachand@iitism.ac.in, Pandeyh@edgehill.ac.uk

**Abstract:** In a WSN, sensor node plays a significant role. Working of sensor node depends upon its battery's life. Replacements of batteries are found infeasible once they are deployed in a remote or unattended area. Plethora of research had been conducted to address this challenge, but they suffer one or the other way. In this paper, a particle swarm optimization (PSO) algorithm integrated with an energy efficient clustering and sink mobility ((PSO-ECSM) is proposed to deal with both cluster head selection problem and sink mobility problem. Extensive computer simulations are conducted to determine the performance of the PSO-ECSM. Five factors such as residual energy, distance, node degree, average energy and energy consumption rate (ECR) are considered for CH selection. An optimum value of these factors is determined through PSO-ECSM algorithm. Further, PSO-ECSM addresses the concern of relaying the data traffic in a multi-hop network by introducing sink mobility. PSO-ECSM's performances are tested against the state-of-the-art algorithms considering five performance metrics (stability period, network, longevity, number of dead nodes against rounds, throughput and network's remaining energy). Statistical tests are conducted to determine the significance of the performance. Simulation results show that the PSO-ECSM improves stability period, half node dead, network lifetime and throughput vis-à-vis ICRPSO by 24.8%, 31.7%, 9.8 %, and 12.2%, respectively.

**Keywords:** Clustering, Energy Consumption Rate (ECR), Energy Efficiency, Optimization, PSO-based CH selection, Sink mobility, Wireless sensor network.

**Abbreviations**

| SYMBOL | PURPOSE |
|---|---|
| PSO-ECSM | Particle Swarm Optimization algorithm integrated with an energy Efficient Clustering and Sink Mobility |
| ECR | Energy Consumption Rate |
| DCH-GA | Dynamic Cluster Head Selection Using Genetic Algorithm |
| GADA-LEACH | Genetic Algorithm based Distance Aware routing -LEACH |
| GABEEC | Genetic Algorithm based energy efficient cluster |
| PSOBS | PSO Based Selection |
| ICRPSO | Inter and Intra Cluster-based Routing using PSO |
| $N_{NORM}$ | Normal Node |
| $N_{ADVN}$ | Advance Node |
| $N_{SUP}$ | Super Node |
| Ʊ | Energy fraction of Super node |
| $\theta$ | Energy fraction of Advanced node |
| $\varpi$ | Energy fraction of advanced node more times compared as normal node |
| $\phi$ | Energy fraction of super node more times compared as normal node |
| $E_O$ | Initial energy of node |
| $E_T$ | Total Energy |
| $FP_S$ | Fitness Parameters |
| $E_{R(i)}$ | Residual energy of $i^{th}$ node |
| E(i) | Initial energy of $i^{th}$ node |
| $D_{(N(i)-s)}$ | Euclidean distance of $i^{th}$ node from the sink |
| $D_{(AVG(N(i)-s))}$ | Average distance Between of $i^{th}$ node and the sink |
| $N_{CL}$ | Number of nodes in the cluster |
| $D_{(N((i)-(j))}$ | Distance Between of $i^{th}$ node and $j^{th}$ node of the cluster |
| RCH | Relay Cluster Head |
| $E_{P(i)}$ | represents the energy value of the $i^{th}$ node in the previous round |

| | |
|---|---|
| $E_{RC(i)}$ | energy consumed in the current round by $i^{th}$ node |
| $E_{R\_CH(i)}$ | residual energy of $i^{th}$ CH node |
| $\varphi$, δ, γ, α, σ, and β | Weight coefficients |
| $N_{CM(i)}$ | `number of nodes in the cluster |
| C(V) | Current sensed value |
| S(V) | Sensed value |
| H(T) | Hard Threshold |
| S(T) | Soft Threshold |
| $E_{tx}$ (z, d) | Energy consumed for transferring the z-bit data at the distance d |
| $E_{elct}$ | Energy consumed for activating the transmitter and receiver circuitry |
| $E_{efs}$ | Free space energy model |
| $E_{amp}$ | Energy consumption for multipath energy model |
| $d_0$ | Threshold Distance |
| $E_{rx}$ | Energy consumed in the reception of z-bit data |
| $E_{da}$ | Energy consumed in the data aggregation of 1-bit data |
| $E_{dx}$ | Energy expenditure during data aggregation |

## 1. Introduction

Wireless sensor networks (WSNs) have proven its applicability across domains which range from military to health care, agricultural to environmental monitoring and many more [1][2]. The working of WSNs largely depends upon sensor nodes. Sensor nodes play an important role for data dissemination from the networks and forwarding it to the sink [1]. The sensor nodes basically comprise of low battery resources, limited range of memory size and computational capabilities at the small magnitude. Basically, the sensor nodes are meant to work in hostile regions, therefore the battery embedded in them is irreplaceable due to this the lifetime of a sensor node is limited. Hence, it is one of the most crucial concerns of WSNs. Literature reveals that utilizing sensor nodes in an energy efficient manner is challenging and it is still an open problem[3]. Bayrakdar proposed energy efficient techniques for terrestrial WSN that made use of the cooperative communication among the sensor nodes [4-6].

Liu [5] suggested cluster-based routing method to preserve the energy of a sensor node. Clustering method helped in grouping the sensor nodes. Each cluster of sensor nodes has assigned one cluster head (CH). CH works as follows: "*collects data from the cluster members and forwards it to the sink*". Tanwar et al. [7] highlighted that clustering is an effective method for energy balancing in the WSNs. Three key merits of clustering method are presented: (a) The communication bandwidth is conserved as the communication of cluster nodes is limited to the CH only. Therefore, the exchange of redundant messages among the nodes is highly avoided; (b) Energy of the sensor nodes is saved in a way that the CH aggregates the collected data therefore, the sink only receives the meaningful data therefore, the energy of the nodes is saved; and (c) Scalability of the network is enhanced tremendously as CHs maintain the local route setup for the other CHs [8].

The research proposed is this paper is based on two rational aspects: (a) How to improve and optimize CH's selection; and (b) How address challenges of sink mobility (a solution to the hot-spot problem). The selection of CH is a crucial optimization problem. It is considered as a NP-hard problem. There are different parameters that must be taken into consideration for the selection of CH namely, residual energy, distance factor, node degree and many more [9]. The thing to be noticed here is the tradeoff between different parameters, e.g. we want node energy to be high but distance to be low from the sink. Therefore, an optimization technique is required to select CH by integrating the important parameters in a designed fitness function [10]. The next problem arises when the network area is large, and the clusters follow multi-hop communication to reduce the energy consumption. Whilst doing so, the clusters located nearer to the sink get burdened with heavy traffic data. Such situation is termed as hot-spot problem [11]. Due to this, eventually sink gets isolated from the network. To resolve this concern, sink mobility has proved to be promising approach [12]. It helps in the distribution of energy around the sink in the most effective manner that balances the load in the network and extends the network lifetime. It is noteworthy to mention that carriers for the mobile sink can be public transport that follows a predefined cyclic schedule to cover a target area. The real time scenario for mobile sink can be mapped to the application of battlefield surveillance. Smart homes, traffic monitoring, pollution control and practicing different safety measures are some of the applications where mobile sinks can be effectively employed [13][9].

While considering the above concerns, we have proposed a novel optimal solution using PSO for CH selection and address the challenges of sink mobility so that an efficient network performance can be achieved. The main contribution of this paper can be stated as follows:

a) Firstly, we propose a novel PSO-ECSM (PSO-based Energy Efficient Clustering and Sink Mobility) algorithm for optimizing CH selection process. Five factors such as energy, distance, node degree, average energy and Energy Consumption Rate (ECR) are considered for the CH selection.
b) Secondly, we showed that the proposed PSO-ECSM algorithm can address challenges of the sink mobility. Challenges of sink mobility are addressed using PSO that considers the lower energy CH and nearest CH from the moving sink and size of the cluster.
c) Third, we comprehensively described the operational steps of the PSO-ECSM and how fitness function for the network is determined.
d) Finally, extensive computer simulations are performed to determine the performance of the PSO-ECSM. Results are compared with the state-of-the-art algorithms. Statistical tests are also conducted to know the statistical significance. We noticed that the proposed PCO-ECSM outperformed other algorithms.

The rest of the manuscript is organized as follows. Section 2 presents related work; Section 3 shows the system framework of the PSO-ECSM; Section 4 outlined the simulation model, results and analysis; Section 5 presents the conclusions.

## 2. Related work

This section shade light on the existing works. Two- fold study is presented: (a) first, we present different routing techniques proposed for CH selection and highlighted reported their limitations; and (b) second, a deep insight to the routing techniques adopting the various meta heuristic including PSO algorithm are highlighted. Study of related work is organized in tabulated form.

Table 1 shows the existing research on CH selection for heterogeneous WSN whilst Table 2 highlight those state-of-the-art researches where PSO algorithm was implemented for routing problems to enhance the network performance.

### 2.1 CH Selection using Heterogeneous Algorithms

The clustering method started from the LEACH [14] which had distributive approach to select CH and selected CH randomly. However, the clustering method improved the network performance from the pre-existing algorithms. Since then a plethora of research has produced numerous variants of LEACH. Tyagi et al. [15] presented a systematic advancements of clustering algorithms. In our study, we focus on the CH selection in heterogeneous WSNs, therefore a systematic review conducted by Elhoseny et al. [16] is considered.

SEP algorithm reported the first heterogeneous routing algorithm that worked upon two levels of energy heterogeneous nodes i.e., normal and advanced nodes [17]. The CH selection was on the weighted probability and not considering various significant parameters like distance, node density, etc. Then, DEEC [18] and DDEEC [19] were reported which worked upon the CH selection considering the residual energy and avoiding the penalization of high energy nodes. EEHC [20] proposed CH selection algorithm for a network employing three level heterogeneous nodes. As there was no provision to avoid penalization of nodes, its network performance was not improved optimally. EDDEEC [21] extenuated this concern by proposing energy threshold concept for the CH selection similar to the DDEEC that performed at two levels. With the advancement in the heterogeneity levels, BEENISH [22] was introduced that had energy heterogeneous nodes at four levels. Akbar et al. [23] proposed IBEENISH that improved the BEENISH in a way that the penalization is avoided at the four-level energy heterogeneous nodes. With the same intent to improve the stability period and network lifetime, Paola et al. [24] proposed P-SEP that selected CH randomly rather than giving any priority to the high energy nodes. These algorithms worked for the single hop communication and the sink placement for them is done inside the network. However, the algorithms DRESEP [25], SEECP [26] and TEDRP [27] worked for dual hop communication. The aforementioned algorithms suffered from the Hot-Spot problem as no provision was made to overcome the burdening of relaying nodes. Verma et al. proposed various methods for CH selection using optimized and non-optimized methods [28]. The authors in [29] used the GA-based optimization for the heterogeneous network. Quality of Service (QoS) were handled to improve the network performance with the use of multiple data sinks in the network [30].

Although, the state-of-art techniques have left no stone unturned in enhancing the network performance, but it is observed that CH selection is NP-Hard (Non-Polynomial Hard) and acquiring the optimal network performance is one of the daunting tasks. So, there is desideratum for some metaheuristic method that can accommodate essential parameters for the process of optimization.

**Table 1**. Classification of various algorithms based on CH selection method and research gaps.

| Reference No. | Name of Algorithms | Heterogeneity level | Reactive/Proactive | Mode of Communication | Hot-spot Problem | No. of CH fixed | CH selection based on | | | | | | Research Gap |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Initial Energy | Residual Energy | Total Energy | Average Energy | Distance | Node Density | |
| [17] | SEP | 2 | Pro-active | single hop | No | × | ✓ | × | × | × | × | × | Not suitable for multi-level |
| [18] | DEEC | 2 | Pro-active | single hop | No | × | × | ✓ | × | ✓ | × | × | Penalization of high energy nodes |
| [20] | EEHC | 3 | Pro-active | single hop | No | × | × | × | × | × | × | × | Energy of nodes is not considered for CH selection |
| [19] | DDEEC | 2 | Pro-active | single hop | No | × | × | ✓ | × | ✓ | × | × | Not suitable for multi-level heterogeneity |
| [21] | EDDEEC | 3 | Pro-active | single hop | No | × | × | ✓ | × | ✓ | × | × | Not suitable for multi-level heterogeneity |
| [22] | BEENISH | 4 | reactive | single hop | No | × | × | ✓ | × | ✓ | × | × | Penalization of high energy nodes |
| [32] | TSEP | 3 | reactive | single hop | No | × | × | × | × | × | × | × | Energy factor not included for CH selection |
| [25] | DRESEP | 3 | reactive | dual hop | Yes | × | × | ✓ | × | × | ✓ | × | Pre-fixed circular radius in random deployment scenario |
| [26] | SEECP | 3 | reactive | dual hop | Yes | ✓ | × | ✓ | × | ✓ | × | × | In-efficient selection of circular radius for dual hop comm. |
| [33] | P-SEP | 2 | Pro-active | single hop | No | × | × | × | × | × | × | × | Selects CH randomly, No preference to advanced node |

**Table 2.** Classification of the state-of-art routing techniques employing PSO to optimize and enhance the network performance.

| Reference No. | Optimization Technique | Targeted attributes | Fitness parameters integrated in Fitness Function | Research Gap |
|---|---|---|---|---|
| [34] | PSO-MSB | Network lifetime, data delivery and energy consumption | Distance between the nodes and sink | → Considering only distance factor in fitness function makes it energy inefficient |
| [35] | Immune orthogonal learning particle swarm optimization algorithm (IOLPSOA) | Network Lifetime, Communication overhead, Repairing routing topology | Energy consumed on a particular path and by sensor node<br>Delay taken by node and Distance of nodes with the edges | → Non-scalable due to lack of clustering<br>→ Too many transmissions will consume energy<br>→ Computational complexity is high |
| [36] | PSO-Semi Distributed (PSO-SD) | Average energy consumption, PSO based number of CH selection, Average number of packets transmission, Network Lifetime | Residual energy,<br>intra-cluster distance, node degree and head count of the probable cluster heads. | → Network area is large, single hop communication will consume energy<br>→ Not significant improvement in network lifetime |
| [37] | PSO | Network life, energy consumption, dead sensor nodes and delivery of total data packets to the base station. | Lifetime of CHs, Average distance of nodes from the CH in cluster | → CH selection only considers the energy and distance factors.<br>→ Hot-spot problem exists as this scheme works for more than one hop among nodes. |
| [38] | E-OEERP (Enhanced Optimized Energy Efficient Routing Algorithms) | Energy consumption, Throughput, Network Lifetime and Packet Delivery Ratio | Ration of Distance between the CH and the cluster member nodes to the number of cluster nodes, Energy of CH to the energy of nodes | → CH selection considers PSO for CH selection and GSA for selecting the best route, however the selection the fitness parameters are not energy efficient. |
| [39] | Fuzzy Clustering and PSO (FCPSO) | Network Lifetime, node mortality rate and Throughput | Distance of a node from the CH and from the CH to the sink, energy consumption encountered for same | → Inefficient parameters for the fitness function.<br>→ Load imbalance is existing<br>→ CH selection is not significantly discussed |
| [40] | Endocrine Cooperative PSO Algorithm (ECPSOA) | Network Lifetime and minimization of energy consumption | Sink mobility is decided with remaining energy of node, distance factor, Energy consumption of node, communication delay function of node | → Due to non-clustering approach, the number of transmissions is very high that leads to energy consumption<br>→ The multi hop communication results in hot-spot problem. |
| [41] | PSO-ECHS (Energy Efficient Cluster Head Selection based on PSO) | Network Lifetime, Minimization of energy consumption, Throughput | Intra-cluster distance, sink distance and residual energy | → Routing algorithm is not discussed. The involvement of other factors for CH selection can improve the network performance. |
| [42] | Energy Efficient clustering and routing using PSO | Network Lifetime and Throughput for different number of CHs | Average energy, standard deviation of remaining energy, and average energy of path of gateways (CH), | → The proposed algorithm is a centralized one, so the global information is required by the sink |
| [43] | PSO based routing | Network Lifetime | Number of relay nodes, distance to the base station, and relay load factor | → Various important parameters are ignored while implementing the proposed technique. |
| [44] | PSO Based Selection (PSOBS) | Number of Hops, Packet Loss Rate, Standard Deviation, Throughput, Energy consumption | Number of rendezvous points, expected number of RPs, maximum tour length, and length of route passing through random RPs | → Inefficient Clustering, CH selection is not discussed |

### 2.2 Role of meta-heuristic approach in CH selection and sink mobility

Since the development of WSN the primary concern of preserving the battery of sensor nodes has been reported in the various literature work existing so far. Other than the conventional methods that aim for CH selection and optimizing the sink mobility in the WSN, there are methods that opt for various meta-heuristic approach for the same. In this literature study we have given tabular study of each; CH selection as well as the sink mobility in WSN. The work focusses on PSO based sink CH selection and sink mobility as given in Table 2. Furthermore, the tabular studies of other meta-heuristic methods are given for the CH selection as well as for sink mobility in Table 3 and Table 4.
Table 3 shows the optimization process involved in the CH selection, the findings and corresponding research gaps. Furthermore, Table 4 discusses the same but only covering the optimization strategy involving sink mobility. It is worth mentioning that among the various existing meta heuristic algorithms, the state-of-the-art strategies are covered.

**Table 3** Meta-heuristic approach for optimized CH selection in WSN

| Ref. No. | Name of meta-heuristic method | Targeted attribute | Findings | Research Gaps |
|---|---|---|---|---|
| Shankar et al. (2016) [45] | Harmony Search Algorithm and PSO | Number of alive and dead nodes, residual energy and throughput | High search efficiency of HSA and dynamic capability of PSO is utilized. | The proposed protocol is deprived of various eminent factors of CH selection |
| Pitchaimanickam and Murugaboopathi (2019) [46] | Hybrid of firefly and PSO | Network Lifetime, energy consumption reduction | The functionality of PSO is utilized in the firefly algorithm which has not been applied toward the optimal cluster head selection problem | Only energy and distance are considered for CH selection |
| Vijayalakshmi and Anandan (2019) [47] | Tabu and PSO | Number of clusters formed, percentage of nodes alive, reduction of packet loss, and average end to end delay | The proposed method helps in reducing the packet loss efficiently | In efficient selection with no explanation of parameters for CH selection |
| Chandirasekaran and Jayabarathi (2019) [48] | Cat Swarm Optimization | Number of alive nodes | Received signal strength, residual battery voltage and intra cluster distance considered for CH selection | The performance evaluation of the proposed work is done against the traditional LEACH protocol |
| John and Rodrigues (2019) [49] | Taylor Crow Optimization (Taylor Series and Crow Search Algorithm) | Number of alive nodes, Normalized node energy and Normalized network energy | The factors considered are distance between the nodes in the cluster, energy of the nodes, traffic density of the cluster, and the delay in transmitting the data packets | The proposed algorithm is computationally complex |
| Alghamdi (2020) [50] | Dragon fly and Firefly | Number of alive nodes Normalized energy Convergence | Energy, delay, distance and security are considered | The CH selection could be further improved by consideration of additional factors |

**Table** 4 Meta-heuristic approach for optimized sink mobility in WSN

| Ref. No. | Name of meta-heuristic method | Targeted attribute | Findings | Research Gaps |
|---|---|---|---|---|
| Gupta and Saha [51] | Artificial Bee Colony and Differential Evolution | Average energy consumption Total residual energy Network Lifetime | average energy, intra-cluster distance and delay parameters Mobile sink is re-localized within a cluster | The parameters are not sufficient enough to acquire optimal performance of the network |
| Krishnan et al. (2019) [52] | Ant Colony Optimization for CH selection and Traveling salesman problem is used for mobility | Network Lifetime, Dead sensor nodes | Residual energy, distance and CH degree are considered for CH selection | The sink mobility is not optimized with efficient parameters |

| | | | | |
|---|---|---|---|---|
| Vijayashree and Dhas [53] | Artificial Bee Colony | Energy consumption Transmission Delay Network connectivity | Number and the location of CHs were considered Random walk is used for the sink mobility | CH selection is done only based on the energy Use of four mobile sink makes the network more computationally complex |
| Zhang et al. (2019) [54] | Ant Colony Optimization | Average delay, Energy consumption, average hops and Lifetime | Density of nodes, relative residual energy, and the degree of uniformity of distribution are considered for weighting the rendezvous nodes | The selection of CH is not optimized therefore, the data collection from the nodes becomes inefficient |
| Wang et al. (2020) [55] | Elite hybrid optimization (PSO, difference operator of differential algorithm and ACO) | Packet reception rate, Network Lifetime | The elite portion of three optimization algorithms is used to find the sink optimized path for movement | The proposed algorithm is incapable of discovering the feasible solutions when the area is large. |

This is because most classical optimization methods are based on a limited number of standard forms, which means that they have to comply with the particular structures of objective functions and constraints. However, in realistic scenarios it is often impossible to accurately characterize the physical problem with an ideal standard-form optimization problem model. Additionally, many complicated factors, such as a large number of integer variables, non-linearity, and so forth may occur. Both of them can make the realistic problems hard to solve. Therefore, the classical mathematical programming-based optimization methods may not be suitable for solving the MOPs encountered in real-world WSNs. Over the most recent decade, metaheuristics have made substantial progress in approximate search methods for solving complex optimization problems. A metaheuristic technique guides a subordinate heuristic using concepts typically derived from the biological, chemical, physical and even social sciences, as well as from artificial intelligence, to improve the optimization performance. Compared to mathematical programming-based methods, metaheuristics-based optimization algorithms are relatively insensitive to the specific mathematical form of the optimization problems. However, the higher the degree of accuracy required, the higher the computational cost becomes. So far, the field of metaheuristics-based optimization algorithms has been mostly constituted by the family of evolutionary algorithms

### 2.2.1 PSO for CH selection

This section focuses on PSO based CH selection. PSO based algorithms are discussed in Table 2. The flow chart for PSO algorithm is depicted in Figure 1.

The numerous particles are initialized which define the solution. Furthermore, every solution is an array of number of the nodes in the cluster. According to the PSO, these particles are checked for their fitness values and accordingly the personal best value for the position of particle is computed. Thereafter, the global best value is determined among all the particles. With the help of these values, the particle's position and velocity is upgraded. The whole process is repeated till it further reaches to the termination due to the completion of total iterations or the stopping criteria is achieved.

Latiff et al. [34] proposed PSO-MSB targeting different attributes namely, network lifetime, data delivery and energy consumption are considered in which the distance between the node and sink is considered. Hu et al. [35] proposed immune orthogonal learning particle swarm optimization algorithm (IOLPSOA). However, the long-haul transmission in this method consumes a lot of energy. The various other algorithms are discussed in the Table 2. The key points inferenced from algorithms reported in Table 2 are states as follow: (a) Many of the algorithms did not consider the energy efficient CH selection while implementing the PSO algorithms for the sink mobility; (b) It is seen that numerous algorithms did not target the hot-spot problem which could save the energy and enhance the network lifetime; and (c) The computational complexity is another factor that makes the existing algorithms non-suitable for various real time applications.

PSOBS algorithm was proposed for the selection of rendezvous points using the PSO optimization algorithm where the sink can be moved [40]. The weight value was assigned to all the nodes in the network based on the data packets it receives from the other sensor nodes. The energy consumption of various nodes was reduced due. However, PSOBS suffers from the limitation that it doesn't select CH efficiently. Gharaei et al. [56] presented ICRPSO (Inter and intra cluster-based Routing using PSO). In ICRPSO, inter and intra cluster movement of mobile sink for clustering by using PSO was discussed. ICRPSO suffered from the following drawbacks: (a) ignoring the data traffic due to spiral motion of mobile sink; and (b) random movement of sink inside the cluster had increased energy consumption.

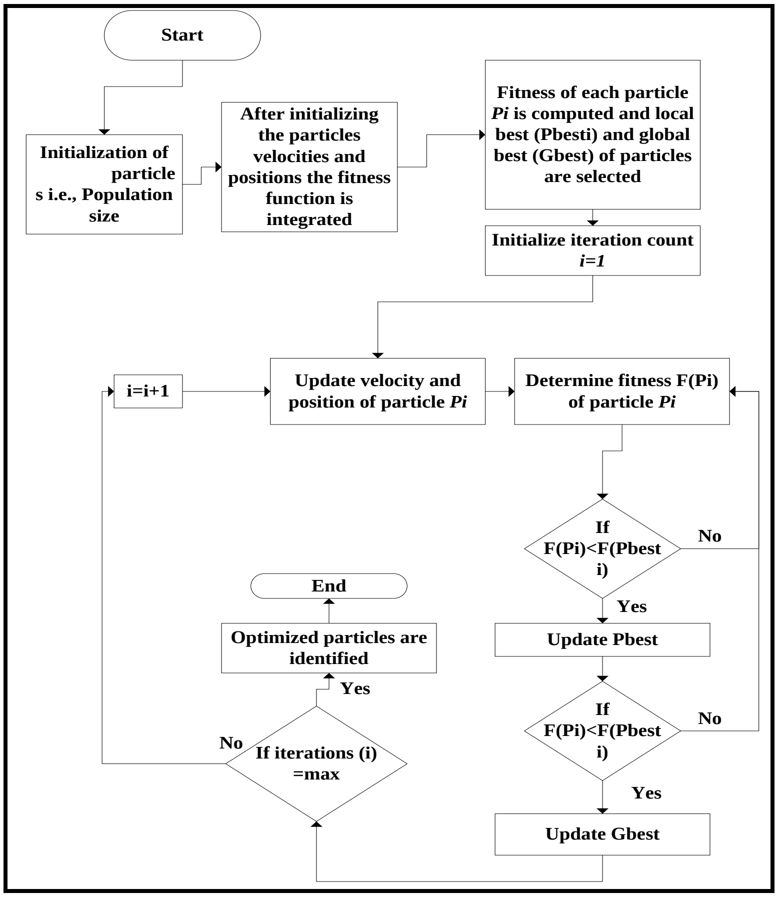

**Figure 1**. Flow chart for PSO optimization process.

Therefore, rather considering two sink, single sink can be employed so that a cost-effective network can be developed. Then, the PSO algorithm can be implemented to optimize the sink mobility.

## 3. The system framework of PSO-ECSM

This section presents the description of the proposed system framework. The heterogeneous model of the PSO-ECSM and its operation involving different functioning steps of PSO are discussed in a comprehensive manner. For the proposed framework the optimization objectives, constraints and decision variables are stated as follow.

a) Optimization objectives

- It is aimed to propose a fitness function that integrates the essential parameters for the sink mobility.
- It is done so to achieve the enhanced network performance.

b) The various constraints are included the limited battery resources of the sensor nodes.

c) The decision variables in the proposed work are considered from the low level of the battery of the sensor nodes to the highest energy stored in the nodes.

### 3.1 Heterogeneous Model of PSO-ECSM

PSO-ECSM utilizes the heterogeneous model for its operation. The nodes are energy heterogeneous and three level of energy heterogeneity is employed in the network as shown in Figure 2. The number of normal, advanced and super nodes used in the network are represented as$N_{NORM}$, $N_{ADVN}$ and$N_{SUP}$as given eq. (1-9). The quantity of these high-energy nodes i.e., advanced and super nodes fraction are represented byƱ and ɵrespectively.

$$\mathrm{N_{SUP}} = \mathrm{n} \times \text{Ʊ} \tag{1}$$

$$\mathrm{N_{ADVN}} = \mathrm{n} \times \text{ɵ} \tag{2}$$

$$\mathrm{N_{NORM}} = \mathrm{n} \times (1 - \text{Ʊ} - \text{ɵ}) \tag{3}$$

The advanced and super nodes are $\varpi$ and $\phi$ times more in energy as compared to normal nodes, respectively. The symbols $\varpi$ and $\phi$ represent the energy fractions of advanced and super nodes, respectively. The computation of total energy of the network represented by $E_T$is done through the set of eq. (4-9).$E_{SUP}$, $E_{ADVN}$, and$E_{NORM}$represent the energy of super, advanced, and normal nodes, respectively.

$$\mathrm{E_{SUP}} = \mathrm{E_O} \times \mathrm{n} \times (\text{Ʊ} + 1) \tag{4}$$
$$\mathrm{E_{ADVN}} = \mathrm{E_O} \times (1 + \phi) \times \mathrm{n} \times \text{ɵ} \tag{5}$$
$$\mathrm{E_{NORM}} = \mathrm{E_O} \times (1 - \text{Ʊ} - \text{ɵ}) \times \mathrm{n} \tag{6}$$
$$\mathrm{E_T} = \mathrm{E_{ADVN}} + \mathrm{E_{SUP}} + \mathrm{E_{NORM}} \tag{7}$$
$$\mathrm{E_T} = \mathrm{E_O} \times (1 + \phi) \times \mathrm{n} \times \text{ɵ} + \mathrm{E_O} \times (1 + \text{Ʊ}) \times \mathrm{n} + \mathrm{E_O} \times (1 - \text{Ʊ} - \text{ɵ}) \times \mathrm{n} \tag{8}$$
$$\mathrm{E_T} = \mathrm{n} \times \mathrm{E_O} \times (1 + \text{Ʊ} \times \text{ɵ} + \text{Ʊ} \times \phi) \tag{9}$$

The total energy computed above is used further in the process of CH selection while integrating fitness function in the following Subsection.

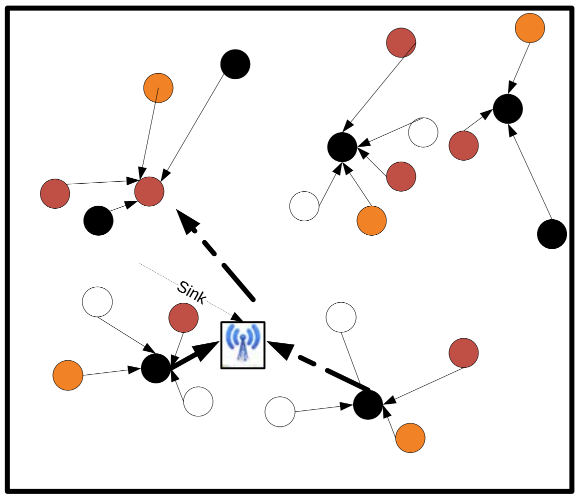


Figure 2 Network demonstration

### 3.2 Working of the PSO-ECSM

The algorithm PSO-ECSM starts working with the validation process where the nodes are represented as the set of particles, which are further evaluated in the form of bit streams. The status of node as a CH is indicated when the bit is '1' otherwise, the node is declared as member node when the bit value is '0'. This process of validation helps in the initialization process in a way that the eligible nodes are taken into consideration for the further steps of optimization.

#### 3.2.1 Initialization

After performing validation, the process of initialization is brought into operation. The certain particles are initialized based on their desirable characteristics. It is analogous to the network parameters comprising network area, number of nodes, sink position in the network and the value of transmitting and receiving energy encountered while data transmission. After performing initialization process, the fitness function is computed

#### 3.2.2 Fitness Function

Fitness function is integration of different performance parameters combined to frame an expression that is to be either maximized or minimized. Fitness function deals with various fitness parameters that decide for the fitness of current individual. The fitness parameters employed in the fitness function as discussed as follows.

#### 3.2.3 Fitness Parameters (FPs)

The FP is computed for its current value depending upon various factors. It is to be noted that more significant the parameter is, more optimized value will be acquired. Here the fitness parameters aim to reduce the energy consumption and rendering the network longevity to the network. The following parameters are taken into consideration while developing fitness function. These parameters are considered for the selection of CH in the network and are discussed as follow.

#### 3.2.3.1 The residual energy of a node

The one of the most prominent factors that is considered while selecting the CH is the residual energy of the node after each round. The reason behind selecting this factor is the rotation of CH based on the residual energy of the node. Rotation of CH is required to bring energy balancing in the network. As the network considered is heterogeneous in nature, the node with maximum energy is favored to be selected as CH. In this parameter, the ratio of residual energy to the total energy is taken into consideration. $\mathrm{FP_{1st}}$(First Fitness Parameter) that pact with energy is defined using equation (10).

$$\mathrm{FP_{1st}} = 1 \Big/ \sum_{i=1}^{N} \left(\frac{\mathrm{E_{R(i)}}}{\mathrm{E_T}}\right) \tag{10}$$

In eq. (10), the summation of ratio of residual energy of $i^{th}$ node represented by $E_{R(i)}$ and total energy denoted by $E_T$ is considered to evaluate the $\mathrm{FP_{1st}}$. The total number of nodes is denoted by *N*. Lower the value of $\mathrm{FP_{1st}}$ for a node, lower will be the chances of selecting it as CH.

#### 3.2.3.2 The average energy of a node

Another energy factor to be considered for the selection of CH is the average energy of a node. The average energy of a node is considered because the network is deployed with the energy heterogeneous nodes. Therefore, the high initial energy nodes are preferred for the selection of CH. It is because super nodes sustain for longer duration as compared to the advanced nodes and similarly advanced nodes are preferred over the normal nodes. The second fitness parameter represented by $\mathrm{FP_{2nd}}$,is the average energy of a node and it is normalized to have value between 0 and 1. In eq. (11), the $E_{(i)}$ represent the energy of the $i^{th}$ node and N represent the total number of nodes in the network.

$$\mathrm{FP_{2nd}} = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} \mathrm{E_{(i)}} \tag{11}$$

#### 3.2.3.3 Distance between sink and node

Whenever nodes are communicating among themselves or with the sink, it is the distance factor that decides the energy consumption of a node under communication. The lesser the distance between node and sink, lesser the energy will be consumed by the node. Therefore, the routing strategies or the CH selection takes care of this parameter so that average distance between the sensor nodes and sink could be minimized. The third Fitness Parameter ($\mathrm{FP_{3rd}}$)for designing the fitness function for the CH selection pacts with distance factor and is given by eq. (12).

$$\mathrm{FP_{3rd}} = \sum_{i=1}^{\mathrm{N}} \left(\frac{\mathrm{D_{(N(i)-S)}}}{\mathrm{D_{AVG(N(i)-S)}}}\right) \tag{12}$$

$\mathrm{FP_{3rd}}$ calculates the summation of distance cost incurred for each *i^th^* node where *i* ranges from 1 to *N* (total number of nodes in the network). In eq. (12), $D_{N(i)-S}$ represents the Euclidean distance of $i^{th}$ node from the sink whereas, $\mathrm{D_{AVG(N(i)-S)}}$represents the average distance between $i^{th}$node and the sink. It is observed that lesser the value of $FP\_3rd$, less it will favor the selection of a node as a CH.

#### 3.2.3.4 Number of neighbors surrounded by a node

When the network area is large, the intra cluster communication becomes a dominant entity. The selection of a node as CH if made independent of number of neighboring nodes to that node, it will result in selecting node as CH which is far located with respect to other nodes. Consequently, the CH node will consume more energy in collecting data form the other nodes in a cluster. Therefore, to avoid such selection, the number of neighboring nodes is taken into consideration. Therefore, fourth Fitness Parameter ($\mathrm{FP_{4th}}$)deals with number of neighboring nodes and is defined by the following eq. (13).

$$\mathrm{FP_{4th}} = \left(\frac{\sum_{\mathrm{i=1,j=1}}^{\mathrm{N_{CL}}} \mathrm{D_{(N(i)-N(j))}}}{\mathrm{N_{CL}}}\right) \tag{13}$$

Where, $D_{(N(i)-N(j))}$ represents the distance between $i^{th}$ node and $j^{th}$ node of the cluster. $N_{CL}$ denotes the number of nodes in the cluster. In such a way, $\mathrm{FP_{4th}}$ must be minimized to make it an energy efficient CH selection.

#### 3.2.3.5 Energy Consumption Rate (ECR)

It is a significant factor that decides the rate of energy consumption of a node and becomes the prominent concern for the selection of CH. It is the difference between the initial energy of the node and the remaining energy of the node after first round. Subsequently, as the number of rounds is proceeded, the energy of the node in the previous round becomes its initial energy. Therefore, ECR is computed and compared with the threshold average value of ECR. If the computed value is found to be lower than the threshold average value, the node becomes eligible to become CH otherwise, it does not qualify for that round to be CH. The fifth fitness parameter of ECR is given by eq. (14).

$$\mathrm{FP}_{5\mathrm{th}}(\mathrm{ECR}) = \sum_{i=1}^{N} \left( \left(\mathrm{E}_{p(\mathrm{i})} - \mathrm{E}_{\mathrm{RC(i)}}\right) \big/ \left(\mathrm{E}_{p(\mathrm{i})}\right) \right) \tag{14}$$

Where the value $E_{RC(i)}$ denotes the energy consumed in the current round by $i^{th}$ node and $E_{p(i)}$ represents the energy value of the $i^{th}$ node in the previous round.

To calculate the threshold average value of ECR for a node, the node with the lowest energy of the cluster is taken into consideration. If any node consumes energy with heavy magnitude and gets lower in value as compared to the lowest energy node, it is not taken into consideration for the CH selection.

#### 3.2.3.6 Fitness Function for the network

The fitness function of the network is the integration of different fitness parameters integrated altogether in a single expression given as follows in eq. (15).

$$\mathrm{F} = \frac{1}{\varphi \times \mathrm{FP}_{1\mathrm{st}} + \delta \times \mathrm{FP}_{2\mathrm{nd}} + \gamma \times \mathrm{FP}_{3\mathrm{rd}} + \alpha \times \mathrm{FP}_{4\mathrm{th}} + \sigma \times \mathrm{FP}_{5\mathrm{th}}} \tag{15}$$

The fitness function represented by F in eq. (15) should be minimized to bring the network performance to the optimum value.

In eq. (16), $\varphi$,$\delta$,$\gamma$,$\alpha$, and $\sigma$are the weight coefficients multiplied with corresponding fitness parameters. These factors are evenly weighted such that it follows eq. (16).

$$\varphi + \delta + \gamma + \alpha + \sigma = 1 \tag{16}$$

Therefore, the main objective function defined for the PSO is given by eq. (15) and the PSO based operations are applied to minimize this function for network lifetime and stability period enhancement.

It is imperative to mention that though we have proposed PSO-based algorithm. However, while formulating a fitness function that considers its different fitness parameters single objective function is considered for optimization which is defined by eq. (15).

The whole process of PSO applied in the proposed work is presented in algorithm 1 which is discussed as follows.

**Algorithm 1** PSO based routing proposed work

**Input:** Initialize gateways $\xi = \{g1, g2, \ldots, gm\}$,
*Next_Hop_G(gi )*, $\forall i$, $1 \le i \le m$ and *NP* size.
**Output:** Route $R : \xi \rightarrow \{\xi + gm{+}1\}$
**Step one:**
Set number of particles *Pi* ,$\forall i$, $1 \le i \le NP$ .
**Step two:**
**for** *i* = 1 to *NP* **do**
Determine *Fitness_f(Pi )* /*Using Eq. (15)*/
*Pbesti= Pi*
**end for**
**Step three:**
*Gbest=* {*Pbestk*|*Fitness_f(Pbestk )*
*min(Fitness_f(Pbesti ),* $\forall i$, $1 \le i \le NP$ )}
**Step four:**
**while** (stopping criteria meet) **do**
**for** *i* = 1 to *NP* **do**
Velocity and position of *Pi* are updated.
Determination of *Fitness_f(Pi )*
**if** *Fitness_f(Pi) <Fitness_f(Pbesti )* **then**
*Pbesti= Pi*
**end if**
**if** *Fitness_f(Pi) <Fitness_f(Gbest)* **then**
*Gbest= Pbesti*
**end if**
**end for**
**end while**
**Step five:**
Determination *Next_Hop(gi ),* $\forall i$, $1 \le i \le NP$ , (i.e., route *R*) using
*Gbest*.

30. **Step six: Terminate**

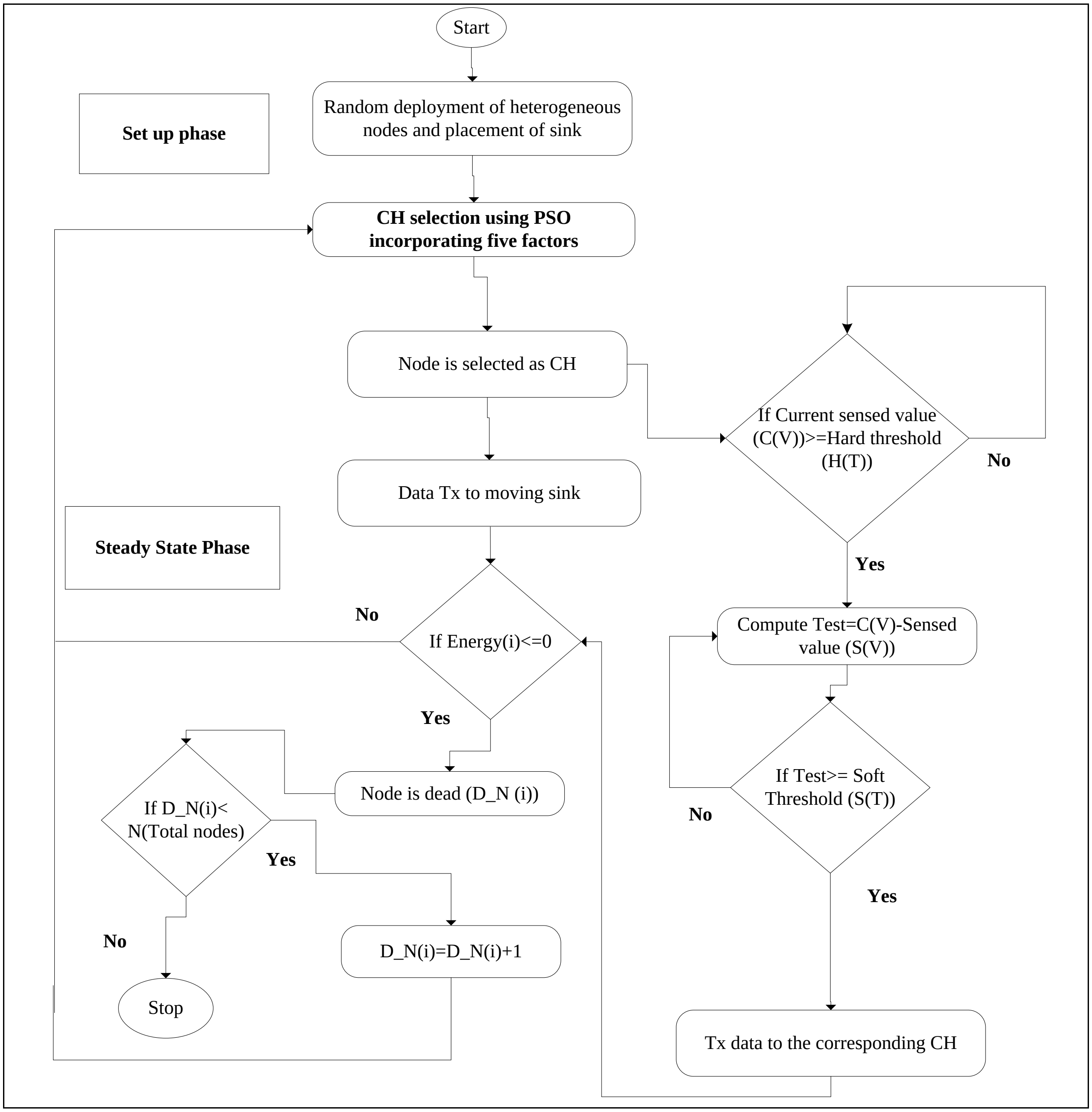


**Figure 3**. Flow chart for working process of PSO-ECSM

Algorithm 1 begins with gateways initialization, the variables *Next_Hop_G(gi )*, *NP*, 'denote the next hop for the data transmission', and the 'population size', respectively. Outputs are represented by the subsequent steps. Steps 1, the number of particles are initialized. In step 2, the fitness of particle is computed i.e., represented by *Fitness_f(Pi )* through eq. (16) and the value is stored in *Pbesti.* In step 3, global best (represented by *Gbest* is computed). In step 4, velocity and position are updated according to the fitness values acquired in previous steps. Furthermore, the fitness of local best (*Fitness_f(Pi))* and global best (*Fitness_f(Gbest))* are compared and the fittest value is selected as global best. Finally, the next hop (*Next_Hop(gi ))* is selected based on the fitness values acquired in the previous steps. These steps are followed till the stopping criteria is reached.

### 3.2.4 Complexity analysis of PSO-ECSM

It becomes imperative to investigate the complexity analysis of proposed algorithm for its feasibility in the real time implementation. As observed from the Algorithm 1 the complexity of the algorithm is found to be O ($r_{max} \times N_p$). Where, $r_{max}$ is denoted as the maximum number of rounds in which network is operated and $N_p$ is the population size respectively.

### 3.3 The Operational steps of PSO-ECSM

The operation of PSO-ECSM is described in flow chart shown in Figure 3. The algorithm is operated under two phases; set up phase and steady state phase which are explained as follow.

a) **Set up phase:** The whole network is structured in this phase i.e., a random deployment of heterogeneous node comprising energy heterogeneity at three levels are deployed in the targeted area. Sink is placed at the middle of the network to collect data from the network forwarding it to the user via internet. Once the node deployment is completed, the clustering of nodes is performed and selection of CH for each cluster is done using PSO. It is to be noted that though the operation of clustering is conventional, but the selection of CH is performed by exploiting five parameters making it promising approach to acquire network lifetime and elongated stability period.
b) **Steady state phase:** Once the setup phase is completed, the PSO-ECSM enters steady state phase where the 'inter and intra cluster' and between the 'CH and sink' communication is performed.

As the PSO-ECSM is a reactive algorithm so it works on the concept of hard and soft threshold as adopted in TSEP algorithm. Initially the data transmission is triggered only when the current sensed value represented by (C(V)) is found to be greater than hard threshold (H(T)) then only data is transmitted by the node to the CH. Furthermore, the transmission in the next round only happens when the difference between the current and previously sensed value is more than the predefined soft threshold. Otherwise, the data transmission is put on hold till the required conditions are satisfied. Once the data transmission is done to the CH then data aggregation is performed at the CH and the useful data is forwarded to the sink. If energy of the node is exhausted completely while doing the data transmission, the node is said to be dead, further counter to the dead nodes is incremented by 1. The same process is repeated till all nodes are dead and the moment when all nodes are dead, the network is said to be dead.

**Lemma 1: PSO-ECSM terminates in fixed iterations $\boldsymbol{It_r = O(1)}$ and renders the optimized CH selection and covers every significant aspect**

***Proof*:** As soon as the set-up phase is completed, the PSO-ECSM enters steady state phase where the energy consumption ticks off. The network runs until all nodes are dead. As the data transmission starts the nodes start depleting their energies and the moment comes when any node completely exhausts its energy. Iterations keep incrementing and it happens until all nodes are dead. The iterations are inversely dependent upon the number of dead nodes. As the number of dead nodes are fixed and constant that makes the operation of PSO-ECSM terminating in fixed iteration.

It is to be noted that the function of PSO-ECSM incorporates various significant factors that considers residual energy and initial energy that favors the CH selection for the nodes which are embedded with more energy at the initial stage and it also considers the available stock of energy for its operation. Furthermore, to abate the energy consumption by the nodes due to distance factor, the distance between the nodes and sink is also taken into consideration.

The methodology for the sink mobility is illustrated in Figure 4 and is discussed through the following steps. The sink mobility is introduced with the PSO optimization technique as the lowest energy CH regions are targeted first to save the energy of those CHs which are about to die. This is computed through the fitness function.

**3.3.1 Fitness parameters:** The fitness parameters for the PSO based sink mobility is given as below.

#### 3.3.1.1 The residual energy of a CH

After the selection of CH from the first algorithm based on PSO, the sink mobility is decided by this parameter. The list for the residual energy of the all selected CHs is formulated in the ascending order. Thereafter, the subsequent movement of sink is done from the lowest energy CH region to that of highest. In this way, the low energy CHs are saved and made to survive long enough to transmit their data to the visiting sink successfully. Hence, the first parameter for the fitness function framed to decide for the sink mobility is given as below.

$\mathrm{FP_{1st}}$ (First Fitness Parameter) that pact with energy is defined as follow.

$$\mathrm{FP_{1st}} = \sum_{i=1}^{\mathrm{N\_CH}} \mathrm{E_{R\_CH(i)}} \qquad (17)$$

In eq. (17), the summation of residual energy of $i^{th}$ CH node is represented by $E_{R\_CH(i)}$ is considered to evaluate the $\mathrm{FP_{1st}}$. The total number of CH nodes is denoted by *N_CH*. Lower the value of $\mathrm{FP_{1st}}$ for a CH node, higher will be the chances of selecting it as a sojourn position to collect data.

#### 3.3.1.2 Distance between sink and CH node

The sink mobility is governed by the distance between the sink and the low energy CH to which sink must move. While considering the low energy, the distance must be taken into consideration so that the sink does not have to move to the far distance to gather data from the nodes. It tends to save time and makes the efficient move of the sink.

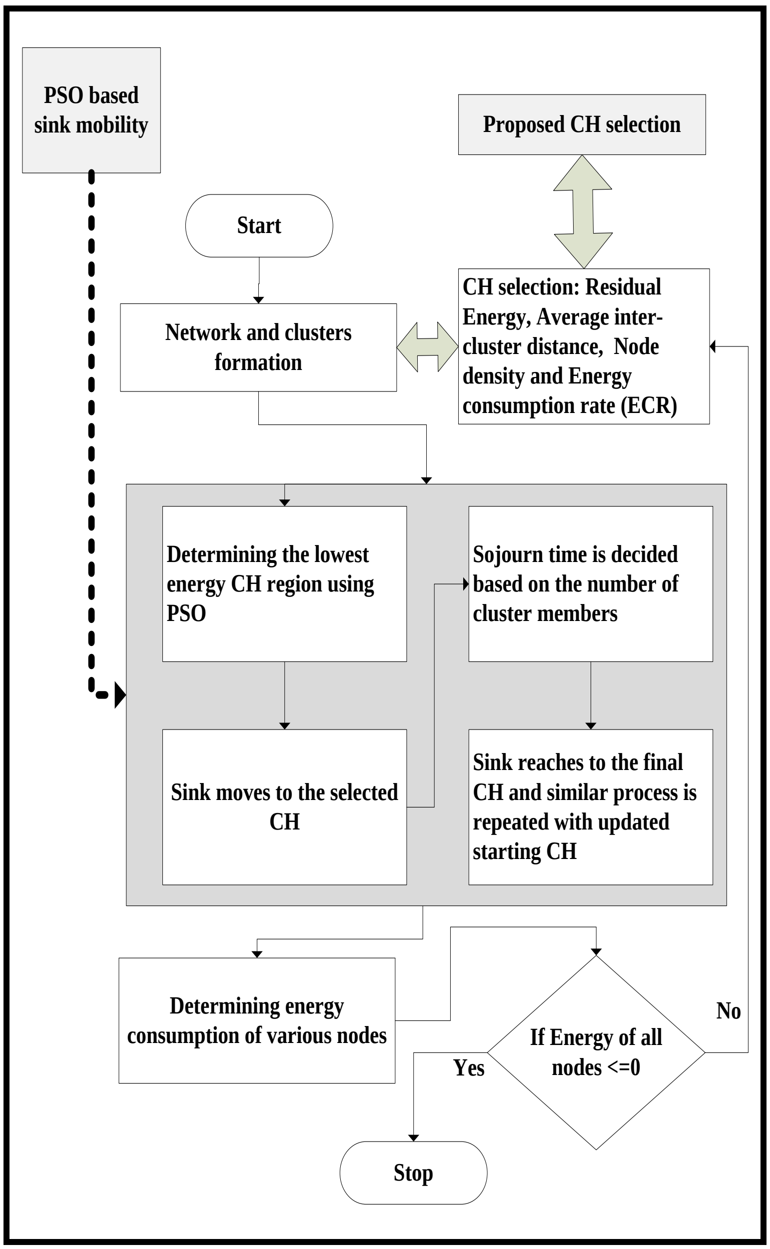


**Figure 4**. The design methodology for the proposed work.

Whenever nodes are communicating among themselves or with the sink, it is the distance factor that decides the energy Therefore, the second Fitness Parameter ($\mathrm{FP_{2nd}}$) for designing the fitness function for the CH selection pacts with distance factor and is given by eq. (18).

$$\mathrm{FP_{2nd}} = \sum_{i=1}^{\mathrm{N\_CH}} \left( \frac{\mathrm{D_{(N\_CH(i)-S)}}}{\mathrm{D_{AVG(N\_CH(i)-S)}}} \right) \tag{18}$$

$\mathrm{FP_{2nd}}$ calculates the summation of distance cost incurred for each $i^{th}$ node where *i* ranges from 1 to *N* (total number of nodes in the network). In eq. (18), $D_{\mathrm{N\_CH(i)-S}}$ represents the Euclidean distance of $i^{th}$ CH node from the sink whereas, $\mathrm{D_{AVG(N_CH(i)-S}}$represents the average distance between $i^{th}$ CH node and the sink. It is observed that lesser the value of $\mathrm{FP_{2nd}}$, more it will favor in selecting the CH node for the sink move towards it.

**3.3.1.3 Size of cluster**

It is quite important to maintain the size of the cluster to the optimum value so that the unbalanced energy consumption could be avoided. When the size of the cluster is high, the energy consumption by that cluster will be high. Ultimately, the movement of sink towards it would save it for a greater number of rounds. The size of the cluster is determined by the number of nodes in the cluster. The third fitness parameter that deals with the size of the cluster is given by eq. (19).

$$\mathrm{FP_{3rd}} = \sum_{i=1}^{\mathrm{N\_CL}} ( N_{\mathrm{CM(i)}}) \tag{19}$$

In eq. (19), $N_{\mathrm{CM(i)}}$ shows the number of nodes in the cluster, and $\mathrm{N\_CL}$ represent the number of clusters in the network.

The size of the cluster should be kept minimum so that the sojourn time could be saved for the other clusters. Therefore, the higher the value of $\mathrm{FP_{3rd}}$, higher will be the chances that the sink moves towards it.

**3.3.1.4 Fitness Function for the network**

As discussed above, the fitness function of the network is the integration of different fitness parameters integrated altogether in a single expression given as follows in eq. (20).

$$\mathrm{F} = \frac{1}{\varphi 1 \times \mathrm{FP_{1st}} + \delta 1 \times \mathrm{FP_{2nd}} + \gamma 1 \times \mathrm{FP_{3rd}}} \tag{20}$$

The fitness function represented by F in eq. (20) should be minimized to bring the network performance to the optimum value.

In eq. (21), $\varphi 1$,$\delta 1$, and$\gamma 1$ are the weight coefficients multiplied with corresponding fitness parameters. These factors are evenly weighted such that it follows eq. (21).

$$\varphi 1 + \delta 1 + \gamma 1 = 1 \tag{21}$$

Therefore, the main objective function defined for the PSO is given by eq. (20) and the PSO based operations are applied to minimize this function for network lifetime and stability period enhancement.

The multi objective PSO-based algorithm is transformed into single objective algorithm which is defined by the single objective function given by eq. (20).

**3.4 System and Network framework of PSO-ECSM**

The heterogeneous model incorporated is discussed in above discussed sections. These nodes suffer from energy consumption when they are involved in data transmission. The radio energy model that decides the amount of energy consumption is discussed in this section.

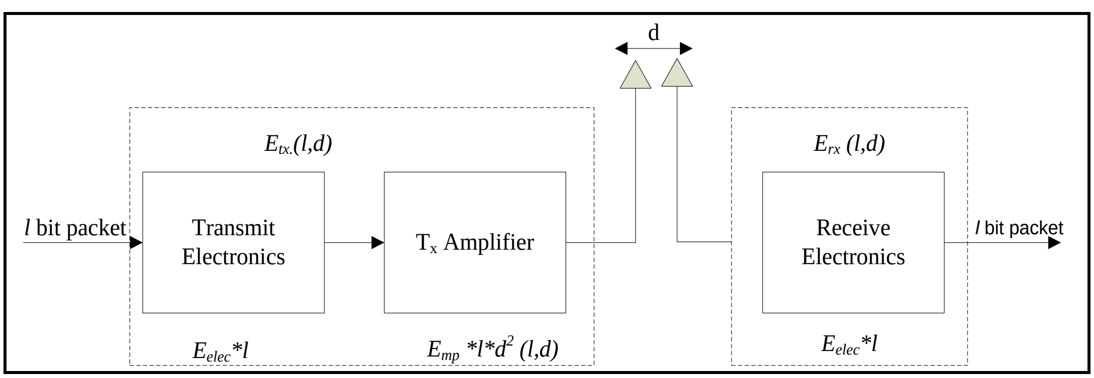


**Figure 5**. Radio Energy Dissipation Model.

**3.4.1 Radio energy model for PSO-ECSM**

The standard radio energy model incorporated in the system framework of PSO-ECSM is shown in Figure 5. The set of eq. (22-26) demonstrates about the equations that are responsible for energy consumptions be it data transmission or data reception. The amount of energy consumption is dependent on the distance between nodes. The energy consumption for transferring the *z* bit data at the distance '*d* 'is denoted by $E_{tx}(z, d)$ and given as follows.

$$E_{tx}(z, d) = z \times E_{elc} + z \times E_{efs} \times d^2 \text{ for } d \leq d_o \quad (22)$$

$$E_{tx}(z, d) = z \times E_{elc} + z \times E_{amp} \times d^4 \text{ for } d > d_o \quad (23)$$

The symbol '$d$' represents the distance between the source and destination nodes or between nodes and sink in eq. (22-23). $E_{elc}$ denotes the energy consumed for activating the transmitter and receiver circuitry. The threshold distance is represented by '$d_o$' and is expressed as in eq. (24).

$$d_o = \sqrt{\frac{E_{efs}}{E_{amp}}} \quad (24)$$

The characteristics of transmitter amplifiers are given by $E_{efs}$ and $E_{amp}$ where $E_{efs}$ is for free space energy model (power loss $d^2$) and $E_{amp}$ denotes the energy consumption for multi path energy model (power loss $d^4$).

The energy consumed while receiving the data per bit is given by eq. (25).

$$E_{rx}(z) = z \times E_{elec} \quad (25)$$

The CH which performs the data aggregation consumes energy as given by the eq. (26).

$$E_{dx}(l) = x \times z \times E_{da} \quad (26)$$

The energy consumed in the reception of *z* bit data, is represented by $E_{rx}$. $E_{da}$ is the energy consumed in the data aggregation of 1-bit data. Moreover, $E_{dx}(z)$ is the energy expenditure during data aggregation of received *z*-bit data of *x* number of data packets.

### 3.4.2 Network model assumptions considered for PSO-ECSM

There are few characteristics of sensor nodes that are taken into consideration while constructing the framework for PSO-ECSM

a) The network is static for the nodes that are immobile in nature and sink moves in the whole network for data collection.
b) The three-level energy heterogeneous nodes are taken into consideration i.e., normal, advanced and super nodes among which super nodes are initiated with highest energy and normal with the lowest one.
c) The sink has no constraint on its energy resources which is quite unlikely to the nodes who get exhausted with energy after certain data transmission rounds.
d) Nodes used are location unaware i.e., no GPS is installed on the node's circuitry.
e) Other factors causing signal attenuation are not taken into consideration.
f) The network area is assumed to be square shaped.
g) The Euclidean distance between nodes is computed by considering the signal strength computed from RSSI (Received Signal Strength Indicator).

### 3.5 Difference in number of parameters in each phase of operation of PSO.

The proposed technique works in the two phases; first being the CH selection which has its own parameters that are proposed to be taken into consideration, and the second pacts with the sink mobility. Evidently, the routing of the data packets from the CHs to the sink has its different dependencies as compared to the CH selection.

When it's about the data transmission to the sink, the factors outside the clusters are considered that involves from whom CH data has to be collected which is decided upon the residual energy of those CH nodes. Moreover, the factors distance between the sink and the CH along with size of cluster is also examined for the sink mobility toward any CH. Hence the CH selection is done based on 5 factors and the sink mobility is accomplished with only 3 parameters.

## 4. Simulation model

The simulation settings decide the environment in which the proposed algorithm is made to operate. The MATLAB software version 2016 is installed on a system with configuration of 2 GB RAM, 1 TB Hard Disk, Intel i3 with CPU operating at 3.07 GHz and Window 10.

### 4.1 Simulation Settings

The network, radio energy model and simulation parameters are given in Table 5. The network is deployed with 100 nodes with three level of energy heterogeneity viz. normal, advanced and super nodes. The energy and amount of nodes fraction is given in Table 5. The parameters of PSO are also given in Table 5 that specifically gives the values for particle size, number of simulations run, and other such parameters are taken into consideration for performing PSO operations for CH selection. The reason behind setting the parameters to the values as mentioned in Table 5 is using the same platform for performance evaluation of the proposed algorithm against the other algorithms. It is done so to have fair comparison of all algorithms on the same platform of network dimensions and parameters. The comparison of the proposed PSO-ECSM is done against the various state of art meta-heuristic approach that basically deals with the Particle Swarm Optimization including ICRPSO, and PSOBS. Furthermore, just to dominate the fact that the improvement is just not countered due to the PSO rather it's the proposed methodology. Therefore, the comparison with the GA based algorithms are also performed that includes algorithms namely, GADA-LEACH, GABEEC and DCH-GA. The tuning of the parameters is done according to the two phases of operation. In first phase, the parameters are tuned for the CH selection. Later, the tuning of the parameters is done for the operation of sink mobility. Multiple simulations have been performed, and the average results are reported. An extensive control parameter tuning is done by Taguchi signal to noise ratio (SNR) method along with orthogonal matrix as done in [58] [59] [60]. Taguchi SNR is a log function of the desired output that serves as an objective function as shown in eq. (27).

$$SNR_i = -10\log\left(\sum_{s=1}^{T_s}\frac{y_s^2}{T_i}\right) \tag{27}$$

Where $i$, $s$, $T_i$ and $y_s$ respectively represent experiment number, trial number, total number of trial for the experiment and number of iterations performed in each trial to get a solution. The simulation scenario for node deployment is given in Table 5.

**Table 5** Simulation parameters.

| **Network Model and PSO Parameters** | **Values** |
|---|---|
| Network Area Size | 100 ×100 m$^2$ |
| Number of Nodes (*N*) | 100 |
| Number of data sinks for PSO-ECSM | 1 |
| Initial energy of nodes (in Joules) $(E_o)$ | 0.5 |
| Energy heterogeneity Node Type | 3-level; normal, intermediate and advanced nodes |
| Energy fraction of intermediate nodes (*β*) and advanced nodes ($\alpha$) | *β=1,* $\alpha$*=2* |
| Number of intermediate nodes (*m*) and advanced nodes fraction ($m_o$) | *m=0.1,* $m_o$*=0.2* |
| Energy required for running transmitter and receiver $E_{elc}$ | 50nJ/bit |
| Threshold distance ($d_o$) | 87m |
| Amplification energy required for smaller distance $d \le d_o$ ($E_{efs}$) | 10pJ/bit/m$^2$ |
| Amplification energy required for larger distance $d > d_o$ ($E_{mp}$) | 0.0013pJ/bit/m$^4$ |
| Energy consumption incurred while data aggregation $(E_{da})$ | 5nJ/bit/signal |
| Data packet size | 2000bits |
| Number of total particles | 30 |
| Initial velocity | 0.0 |
| Initial position | 9.6, 13.6 |
| Number of Simulation run | 20 |

### 4.2 State-of-the-art algorithms for comparison.

For a fair performance validation of the PSO-ECSM, the comparison is done against the metaheuristic techniques employing GA but operating on different fitness function and different routing strategies. The algorithms ICRPSO [56], GADA-LEACH [8], PSOBS [44], GABEEC [57] and DCH-GA [16] are considered for evaluating the performance of the proposed PSO-ECSM.

### 4.3 Simulation Result and Discussion

Multiple simulations have been performed and average results are reported. The results have been analyzed based on the performance metrics. Five different performance metrics (e.g. (a) stability period; (b) network life time; (c) number of dead nodes against rounds; (d) network's remaining energy; and (e) throughput/Number of data packets sent to sink) are used to report the results.

a) **Stability Period:** Results reveals that in PSO-ECSM, the first node is dead after 6301 rounds whereas for ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA, the values of stability period are respectively 5048, 4400, 4370, 4390 and 3608 rounds as shown in Figure 6. It is comprehended that PSO-ECSM ameliorates stability period by 24.8%, 43.2%, 43.5%, 43.5%and 74.6% in comparison to the algorithms ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA respectively. The dominant cause behind such improvement is the incorporation of five fitness parameters that ensure the energy preservation while the data transmission is in process. The distance among the nodes and sink, and within the nodes is effectively decreased. It can be comprehended that PSO-ECSM outperforms the state-of-art algorithms in the context of stability period.

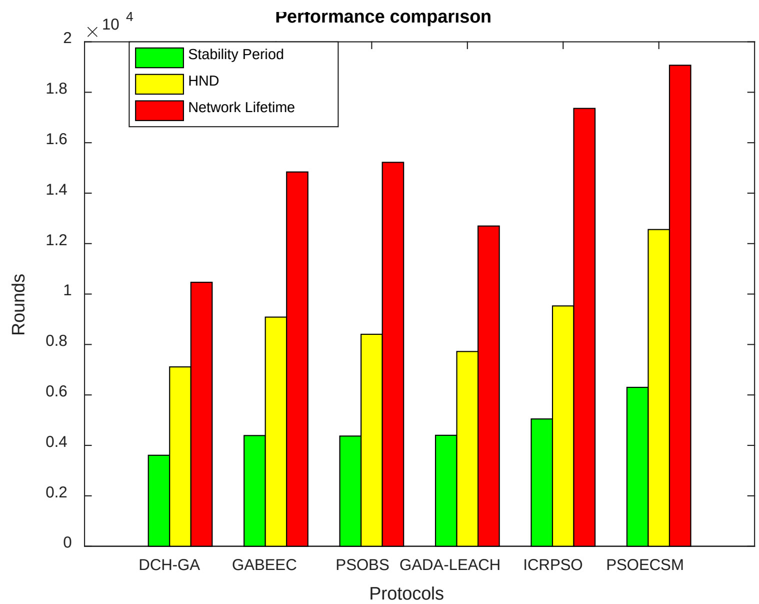


**Figure 6**. Stability Period, HND and Network lifetime comparison of PSO-ECSM with other algorithms.

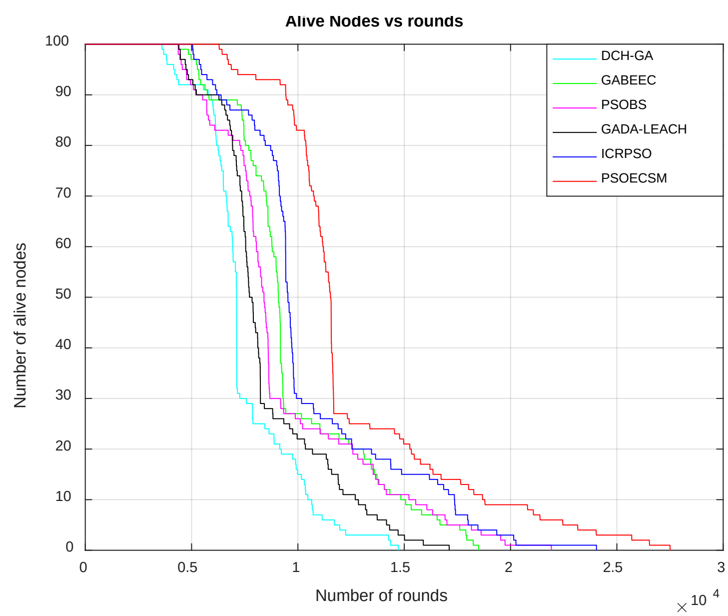


**Figure 7.** Comparison of alive nodes vs rounds of PSO-ECSM with other algorithms.

**b) Network Lifetime:** The network lifetime in PSO-ECSM is achieved at 19071 rounds whereas the network lifetime in case of ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA is observed at 17360, 12697, 15222, 14840 and 10466 rounds, respectively. The percentage improvement in the proposed algorithm is found to be 9.8%, 50.2%, 25.2%, 28.5% and 82.2% as compared to ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA algorithms, respectively as shown in fig 7. The improvement in network lifetime is observed due to the ECR and other factors integrated in fitness function exploited for PSO-ECSM. Furthermore, with high number of neighboring nodes, the average distance between CH and node is decreased comprehensively.

c) **Number of dead nodes against rounds:** the numbers of dead nodes with respect of number of rounds are less in case of PSO-ECSM as compared to the other algorithms. As shown in Figure 8, statistically, the First Node Dead (FND) for PSO-ECSM is 6301 rounds which is 5048, 4400, 4370, 4390, and 3608 rounds in case of ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA, respectively. Half Nodes Dead (HND) for PSO-ECSM is 12558, whereas it is just 9529, 7722, 8405, 9086 and 7112 rounds in case of ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA, algorithms, respectively as shown in Figure 8. Furthermore, as shown in Figure 7, the improvement in Last node dead (LND) i.e., also termed as network lifetime, is also observed in PSO-ECSM as it covers 19071 rounds whereas ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA cover17360, 12697, 15222, 14840 and 10466 rounds, respectively. Such improvement is reported because when the CH selection is optimized under different attributes, more energy preservation is acquired in PSO-ECSM as compared to other algorithms, respectively.

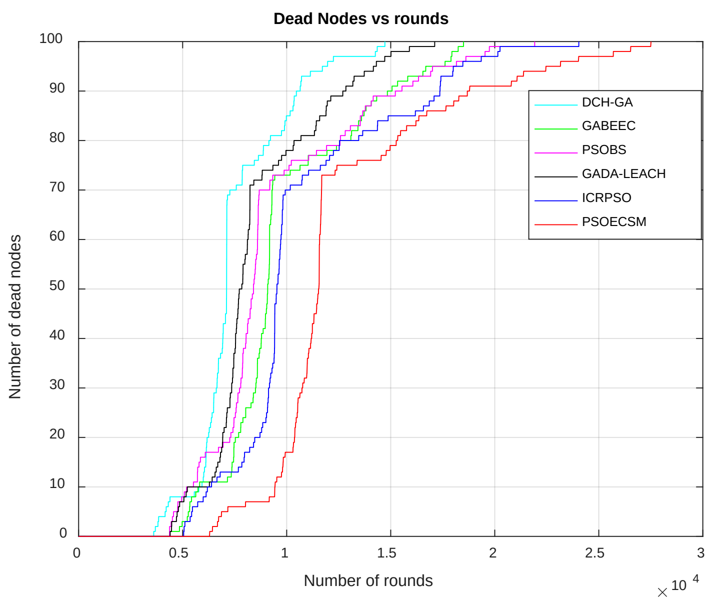


**Figure 8**. Comparison of dead nodes vs rounds of PSO-ECSM with other algorithms.

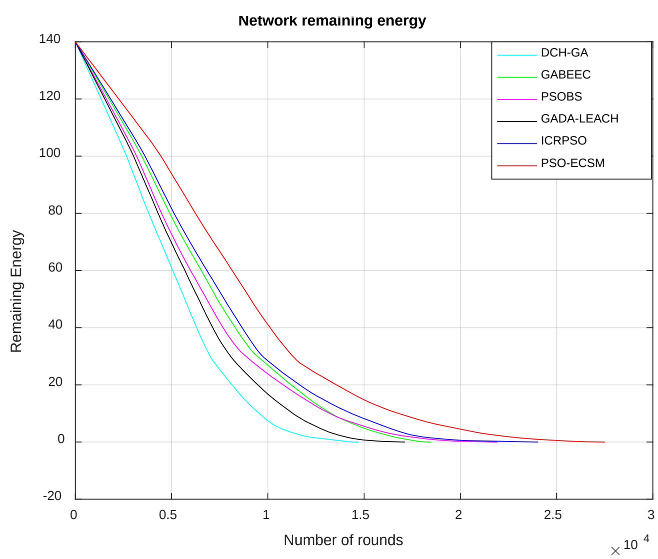


**Figure 9**. Comparison of Network's remaining energy of PSO-ECSM with other algorithms.

**d) Network's remaining energy:** The moment when data transmission is presumed, the network's energy start reducing. It is quite essential to observe the behavior of network's remaining energy with increase in number of rounds. PSO-ECSM performs better as compared to ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA algorithms, respectively in a way that it covers a greater number of rounds while the data transmission is in progress as shown in Figure 9. The energy of a node is preserved in each round due to the minimum energy consumption resulted due to the energy efficient dual hop communication.

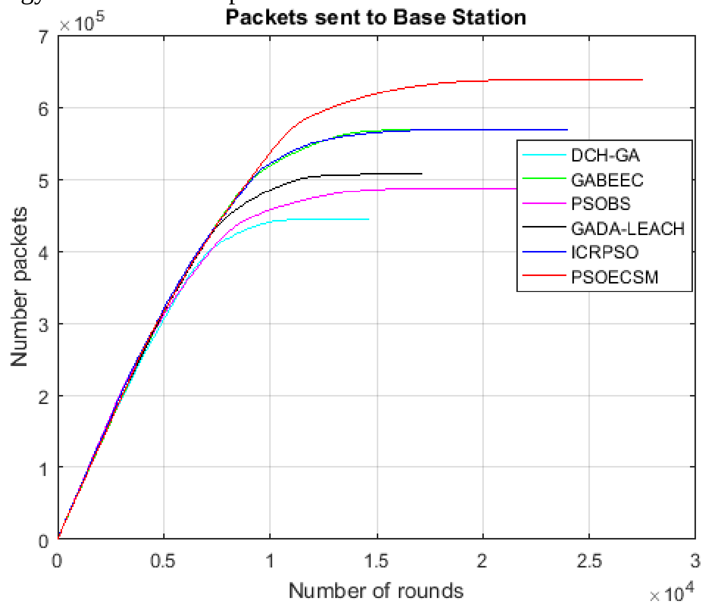


**Figure 10**. Comparison of throughput of PSO-ECSM with other algorithms.

**e) Throughput/Number of data packets sent to sink:** In case of PSO-ECSM, as illustrated in Figure 10, the throughput is enhanced comprehensively as it successfully transmits 637880 data packets whereas ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA transmit 568457, 506203, 486712, 409610 and 308695 data packets, respectively. It is evident from the throughput comparative analysis, PSO-ECSM improves throughput by 12.2%, 26%, 31%, 55.7% and 106.6% as compared to ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA algorithms, respectively. Throughput is enhanced gigantically because with the selection of optimized CH, the data packets are forwarded in the most optimal way leading to the reduction in packet loss during transmission.

### 4.4 Statistical Analysis

Rigorous statistical analysis is performed to determine the significance of the PSO-ECSM. F-test is conducted on the collect samples considering the hypothesis: "*there is no significant difference in the mean of samples at 5% level of confidence*" as presented in equation (28).

$$H_0 : \mu_{DCH-GA} = \mu_{GABEEC} = \mu_{GADA-LEACH} = \mu_{ICRPSO} = \mu_{PSOBS} = \mu_{PSOECSM} \tag{28}$$

$H_A$ : at least one algorithm mean is different than the other.

**Table 6** Descriptive analysis for remaining energy with respective to algorithms.

| Algorithms | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| DCH-GA | 50 | -0.01 | .003 | 0.000 | -0.01 | -0.014 | -0.02 | -0.01 |
| GABEEC | 50 | 2.15 | .026 | 0.003 | 2.14 | 2.16 | 2.11 | 2.20 |
| GADA-LEACH | 50 | 0.15 | .003 | 0.000 | 0.15 | 0.15 | 0.15 | 0.16 |
| ICRPSO | 50 | 4.77 | .035 | 0.004 | 4.76 | 4.78 | 4.71 | 4.83 |
| PSOBS | 50 | 3.01 | .019 | 0.002 | 3.01 | 3.02 | 2.98 | 3.05 |
| **PSOECSM** | 50 | **11.00** | .034 | 0.004 | 10.99 | 11.01 | 10.95 | 11.06 |
| Total | 300 | 3.51 | 3.73 | 0.215 | 3.08 | 3.93 | -0.02 | 11.06 |

**Table 7**. Result of ANOVA test with respect to remaining energy.

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| Between Groups | **(Combined)** | **4174.77** | **5** | **834.95** | **1436395.41** | **0.0** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Linear Term | Contrast | 2770.65 | 1 | 2770.65 | 4766424.24 | 0.0 |
| | | Deviation | 1404.12 | 4 | 351.03 | 603888.20 | 0.0 |
| Within Groups | | | 0.17 | 294 | 0.001 | | |
| Total | | | 4174.94 | 299 | | | |

**Table 8**. Multiple comparison tests (Post hoc test: Tukey HSD and LSD) with respect to remaining energy.

| Test type | (I) Algorithm | (J) Algorithm | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Tukey HSD | DCH-GA | GABEEC | -2.17* | .004 | 0.0 | -2.18 | -2.15 |
| | | GADA-LEACH | -0.16* | .004 | 0.0 | -0.17 | -0.15 |
| | | ICRPSO | -4.78* | .004 | 0.0 | -4.80 | -4.77 |
| | | PSOBS | -3.03* | .004 | 0.0 | -3.04 | -3.01 |
| | | PSOECSM | -11.01* | .004 | 0.0 | -11.03 | -11.00 |
| | GABEEC | DCH-GA | 2.17* | .004 | 0.0 | 2.15 | 2.18 |
| | | GADA-LEACH | 2.00* | .004 | 0.0 | 1.99 | 2.02 |
| | | ICRPSO | -2.61* | .004 | 0.0 | -2.62 | -2.60 |
| | | PSOBS | -0.85* | .004 | 0.0 | -0.87 | -0.84 |
| | | PSOECSM | -8.84* | .004 | 0.0 | -8.85 | -8.83 |
| | GADA-LEACH | DCH-GA | 0.16* | .004 | 0.0 | 0.15 | 0.17 |
| | | GABEEC | -2.00* | .004 | 0.0 | -2.02 | -1.99 |
| | | ICRPSO | -4.62* | .004 | 0.0 | -4.63 | -4.60 |
| | | PSOBS | -2.86* | .004 | 0.0 | -2.87 | -2.85 |
| | | PSOECSM | -10.85* | .004 | 0.0 | -10.86 | -10.83 |
| | ICRPSO | DCH-GA | 4.78* | .004 | 0.0 | 4.77 | 4.80 |
| | | GABEEC | 2.61* | .004 | 0.0 | 2.60 | 2.62 |
| | | GADA-LEACH | 4.62* | .004 | 0.0 | 4.60 | 4.63 |
| | | PSOBS | 1.75* | .004 | 0.0 | 1.74 | 1.76 |
| | | PSOECSM | -6.22* | .004 | 0.0 | -6.24 | -6.21 |
| | PSOBS | DCH-GA | 3.03* | .004 | 0.0 | 3.01 | 3.04 |
| | | GABEEC | 0.85* | .004 | 0.0 | 0.84 | 0.87 |
| | | GADA-LEACH | 2.86* | .004 | 0.0 | 2.85 | 2.87 |
| | | ICRPSO | -1.75* | .004 | 0.0 | -1.76 | -1.74 |
| | | PSOECSM | -7.98* | .004 | 0.0 | -7.99 | -7.97 |
| | **PSOECSM** | **DCH-GA** | **11.01*** | .004 | 0.0 | 11.00 | 11.03 |
| | | **GABEEC** | **8.84*** | .004 | 0.0 | 8.83 | 8.85 |
| | | **GADA-LEACH** | **10.85*** | .004 | 0.0 | 10.83 | 10.86 |
| | | **ICRPSO** | **6.22*** | .004 | 0.0 | 6.21 | 6.24 |
| | | **PSOBS** | **7.98*** | .004 | 0.0 | 7.97 | 7.99 |
| LSD | DCH-GA | GABEEC | -2.17* | .004 | 0.0 | -2.18 | -2.16 |
| | | GADA-LEACH | -0.16* | .004 | 0.0 | -0.17 | -0.15 |
| | | ICRPSO | -4.78* | .004 | 0.0 | -4.79 | -4.77 |
| | | PSOBS | -3.03* | .004 | 0.0 | -3.04 | -3.02 |
| | | PSOECSM | -11.01* | .004 | 0.0 | -11.02 | -11.00 |
| | GABEEC | DCH-GA | 2.17* | .004 | 0.0 | 2.16 | 2.18 |
| | | GADA-LEACH | 2.00* | .004 | 0.0 | 1.99 | 2.01 |
| | | ICRPSO | -2.61* | .004 | 0.0 | -2.62 | -2.60 |
| | | PSOBS | -0.85* | .004 | 0.0 | -0.86 | -0.84 |
| | | PSOECSM | -8.84* | .004 | 0.0 | -8.85 | -8.83 |
| | GADA-LEACH | DCH-GA | 0.16* | .004 | 0.0 | 0.15 | 0.17 |
| | | GABEEC | -2.00* | .004 | 0.0 | -2.01 | -1.99 |
| | | ICRPSO | -4.62* | .004 | 0.0 | -4.63 | -4.61 |
| | | PSOBS | -2.86* | .004 | 0.0 | -2.87 | -2.85 |
| | | PSOECSM | -10.85* | .004 | 0.0 | -10.85 | -10.84 |
| | ICRPSO | DCH-GA | 4.78* | .004 | 0.0 | 4.77 | 4.79 |
| | | GABEEC | 2.61* | .004 | 0.0 | 2.60 | 2.62 |
| | | GADA-LEACH | 4.62* | .004 | 0.0 | 4.61 | 4.63 |
| | | PSOBS | 1.75* | .004 | 0.0 | 1.74 | 1.76 |
| | | PSOECSM | -6.22* | .004 | 0.0 | -6.23 | -6.22 |
| | PSOBS | DCH-GA | 3.03* | .004 | 0.0 | 3.02 | 3.04 |
| | | GABEEC | 0.85* | .004 | 0.0 | 0.84 | 0.86 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | GADA-LEACH | 2.86* | .004 | 0.0 | 2.85 | 2.87 |
| | | ICRPSO | -1.75* | .004 | 0.0 | -1.76 | -1.74 |
| | | PSOECSM | -7.98* | .004 | 0.0 | -7.99 | -7.97 |
| | **PSOECSM** | **DCH-GA** | **11.01*** | .004 | 0.0 | 11.00 | 11.02 |
| | | **GABEEC** | **8.84*** | .004 | 0.0 | 8.83 | 8.85 |
| | | **GADA-LEACH** | **10.85*** | .004 | 0.0 | 10.84 | 10.85 |
| | | **ICRPSO** | **6.22*** | .004 | 0.0 | 6.22 | 6.23 |
| | | **PSOBS** | **7.98*** | .004 | 0.0 | 7.97 | 7.99 |
| *. The mean difference is significant at the 0.05 level. | | | | | | | |

**Table 9**. Homogeneity test based on the mean remaining energy value.

| | | | Subset for alpha = 0.05 | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Algorithm | N | 1 | 2 | 3 | 4 | 5 | 6 |
| Tukey HSD[a] | DCH-GA | 50 | -.0150 | | | | | |
| | GADA-LEACH | 50 | | .1510 | | | | |
| | GABEEC | 50 | | | 2.1572 | | | |
| | PSOBS | 50 | | | | 3.0165 | | |
| | ICRPSO | 50 | | | | | 4.7718 | |
| | **PSOECSM** | **50** | | | | | | **11.0014** |
| | Sig. | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 50.000.

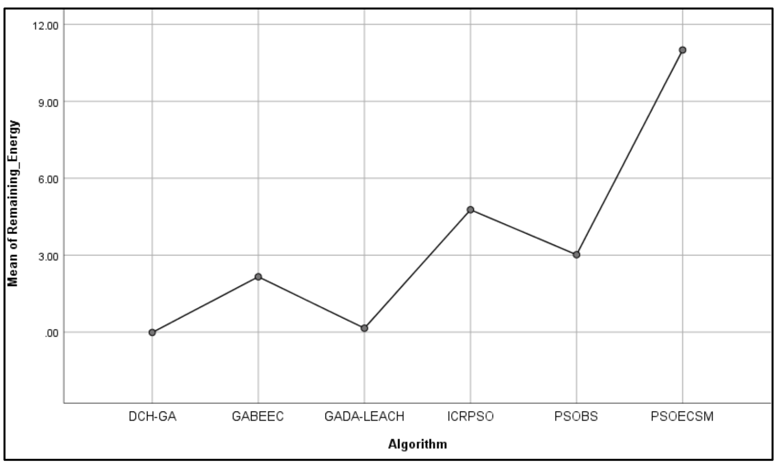


**Figure 11.** Estimated marginal mean plot for remaining energy with respect to each algorithm.

**Table 10**. Descriptive analysis for throughput with respective to algorithms.

| Algorithms | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| DCH_GA | 50 | 443841.26 | .44 | .06 | 443841.13 | 443841.38 | 443841.00 | 443842.00 |
| GABEEC | 50 | 567636.00 | .00 | .00 | 567636.00 | 567636.00 | 567636.00 | 567636.00 |
| GADA-LEACH | 50 | 506197.36 | .48 | .06 | 506197.22 | 506197.49 | 506197.00 | 506198.00 |
| ICRPSO | 50 | 567090.60 | 14.94 | 2.11 | 567086.35 | 567094.84 | 567057.00 | 567107.00 |
| PSOBS | 50 | 486241.36 | 7.10 | 1.00 | 486239.34 | 486243.37 | 486238.00 | 486259.00 |
| **PSOECSM** | 50 | **626778.28** | 60.00 | 8.48 | 626761.22 | 626795.33 | 626665.00 | 626859.00 |
| Total | 300 | 532964.14 | 60668.26 | 3502.68 | 526071.10 | 539857.17 | 443841.00 | 626859.00 |

**Table 11**. Result of ANOVA test with respect to throughput.

| | | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|---|
| Between Groups | **(Combined)** | | **1100510621696.09** | **5** | **220102124339.21** | **340778629.400** | **0.0** |
| | Linear Term | Contrast | 382098426862.24 | 1 | 382098426862.24 | 591593464.139 | 0.0 |
| | | Deviation | 718412194833.85 | 4 | 179603048708.46 | 278074920.716 | 0.0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Quadratic Term | Contrast | 21920650.14 | 1 | 21920650.14 | 33939.196 | 0.0 |
| | | Deviation | 718390274183.70 | 3 | 239463424727.90 | 370755247.889 | 0.0 |
| Within Groups | | | 189888.74 | 294 | 645.88 | | |
| Total | | | 1100510811584.83 | 299 | | | |

**Table 12**. Result of multiple comparison tests (Post hoc test: Tukey HSD and LSD) with respect to throughput.

| Test Type | (I) Algorithms | (J) Algorithms | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Tukey HSD | DCH_GA | GABEEC | -123794.74* | 5.08 | 0.0 | -123809.32 | -123780.15 |
| | | GADA-LEACH | -62356.10* | 5.08 | 0.0 | -62370.68 | -62341.51 |
| | | ICRPSO | -123249.34* | 5.08 | 0.0 | -123263.92 | -123234.75 |
| | | PSOBS | -42400.10* | 5.08 | 0.0 | -42414.68 | -42385.51 |
| | | PSOECSM | -182937.02* | 5.08 | 0.0 | -182951.60 | -182922.43 |
| | GABEEC | DCH_GA | 123794.74* | 5.08 | 0.0 | 123780.15 | 123809.32 |
| | | GADA-LEACH | 61438.64* | 5.08 | 0.0 | 61424.05 | 61453.22 |
| | | ICRPSO | 545.40* | 5.08 | 0.0 | 530.81 | 559.98 |
| | | PSOBS | 81394.64* | 5.08 | 0.0 | 81380.05 | 81409.22 |
| | | PSOECSM | -59142.28* | 5.08 | 0.0 | -59156.86 | -59127.69 |
| | GADA-LEACH | DCH_GA | 62356.10* | 5.08 | 0.0 | 62341.51 | 62370.68 |
| | | GABEEC | -61438.64* | 5.08 | 0.0 | -61453.22 | -61424.05 |
| | | ICRPSO | -60893.24* | 5.08 | 0.0 | -60907.82 | -60878.65 |
| | | PSOBS | 19956.00* | 5.08 | 0.0 | 19941.41 | 19970.58 |
| | | PSOECSM | -120580.92* | 5.08 | 0.0 | -120595.50 | -120566.33 |
| | ICRPSO | DCH_GA | 123249.34* | 5.08 | 0.0 | 123234.75 | 123263.92 |
| | | GABEEC | -545.40* | 5.08 | 0.0 | -559.98 | -530.81 |
| | | GADA-LEACH | 60893.24* | 5.08 | 0.0 | 60878.65 | 60907.82 |
| | | PSOBS | 80849.24* | 5.08 | 0.0 | 80834.65 | 80863.82 |
| | | PSOECSM | -59687.68* | 5.08 | 0.0 | -59702.26 | -59673.09 |
| | PSOBS | DCH_GA | 42400.10* | 5.08 | 0.0 | 42385.51 | 42414.68 |
| | | GABEEC | -81394.64* | 5.08 | 0.0 | -81409.22 | -81380.05 |
| | | GADA-LEACH | -19956.00* | 5.08 | 0.0 | -19970.58 | -19941.41 |
| | | ICRPSO | -80849.24* | 5.08 | 0.0 | -80863.82 | -80834.65 |
| | | PSOECSM | -140536.92* | 5.08 | 0.0 | -140551.50 | -140522.33 |
| | **PSOECSM** | **DCH_GA** | **182937.02*** | 5.08 | 0.0 | 182922.43 | 182951.60 |
| | | **GABEEC** | **59142.28*** | 5.08 | 0.0 | 59127.69 | 59156.86 |
| | | **GADA-LEACH** | **120580.92*** | 5.08 | 0.0 | 120566.33 | 120595.50 |
| | | **ICRPSO** | **59687.68*** | 5.08 | 0.0 | 59673.09 | 59702.26 |
| | | **PSOBS** | **140536.92*** | 5.08 | 0.0 | 140522.33 | 140551.50 |
| LSD | DCH_GA | GABEEC | -123794.74* | 5.08 | 0.0 | -123804.74 | -123784.73 |
| | | GADA-LEACH | -62356.10* | 5.08 | 0.0 | -62366.10 | -62346.09 |
| | | ICRPSO | -123249.34* | 5.08 | 0.0 | -123259.34 | -123239.3 |
| | | PSOBS | -42400.10* | 5.08 | 0.0 | -42410.10 | -42390.09 |
| | | PSOECSM | -182937.02* | 5.08 | 0.0 | -182947.02 | -182927.01 |
| | GABEEC | DCH_GA | 123794.74* | 5.08 | 0.0 | 123784.73 | 123804.74 |
| | | GADA-LEACH | 61438.64* | 5.08 | 0.0 | 61428.63 | 61448.64 |
| | | ICRPSO | 545.40* | 5.08 | 0.0 | 535.39 | 555.40 |
| | | PSOBS | 81394.64* | 5.08 | 0.0 | 81384.63 | 81404.64 |
| | | PSOECSM | -59142.28* | 5.08 | 0.0 | -59152.28 | -59132.27 |
| | GADA-LEACH | DCH_GA | 62356.10* | 5.08 | 0.0 | 62346.09 | 62366.10 |
| | | GABEEC | -61438.64* | 5.08 | 0.0 | -61448.64 | -61428.63 |
| | | ICRPSO | -60893.24* | 5.08 | 0.0 | -60903.24 | -60883.23 |
| | | PSOBS | 19956.00* | 5.08 | 0.0 | 19945.99 | 19966.00 |
| | | PSOECSM | -120580.92* | 5.08 | 0.0 | -120590.92 | -120570.91 |
| | ICRPSO | DCH_GA | 123249.34* | 5.08 | 0.0 | 123239.33 | 123259.34 |
| | | GABEEC | -545.40* | 5.08 | 0.0 | -555.40 | -535.39 |
| | | GADA-LEACH | 60893.24* | 5.08 | 0.0 | 60883.23 | 60903.24 |
| | | PSOBS | 80849.24* | 5.08 | 0.0 | 80839.23 | 80859.24 |
| | | PSOECSM | -59687.68* | 5.08 | 0.0 | -59697.68 | -59677.67 |
| | PSOBS | DCH_GA | 42400.10* | 5.08 | 0.0 | 42390.09 | 42410.10 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | GABEEC | -81394.64* | 5.08 | 0.0 | -81404.64 | -81384.63 |
| | | GADA-LEACH | -19956.00* | 5.08 | 0.0 | -19966.00 | -19945.99 |
| | | ICRPSO | -80849.24* | 5.08 | 0.0 | -80859.24 | -80839.23 |
| | | PSOECSM | -140536.92* | 5.08 | 0.0 | -140546.92 | -140526.91 |
| | **PSOECSM** | **DCH_GA** | **182937.02*** | 5.08 | 0.0 | 182927.01 | 182947.02 |
| | | **GABEEC** | **59142.28*** | 5.08 | 0.0 | 59132.27 | 59152.28 |
| | | **GADA-LEACH** | **120580.92*** | 5.08 | 0.0 | 120570.91 | 120590.92 |
| | | **ICRPSO** | **59687.68*** | 5.08 | 0.0 | 59677.67 | 59697.68 |
| | | **PSOBS** | **140536.92*** | 5.08 | 0.0 | 140526.91 | 140546.92 |

*. The mean difference is significant at the 0.05 level.

**Table 13**. Homogeneity test based on the mean throughput value.

| | | | Subset for alpha = 0.05 | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Algorithms | N | 1 | 2 | 3 | 4 | 5 | 6 |
| Tukey HSD[a] | DCH_GA | 50 | 443841.26 | | | | | |
| | PSOBS | 50 | | 486241.36 | | | | |
| | GADA-LEACH | 50 | | | 506197.36 | | | |
| | ICRPSO | 50 | | | | 567090.60 | | |
| | GABEEC | 50 | | | | | 567636.00 | |
| | **PSOECSM** | **50** | | | | | | **626778.28** |
| | Sig. | | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 50.000.

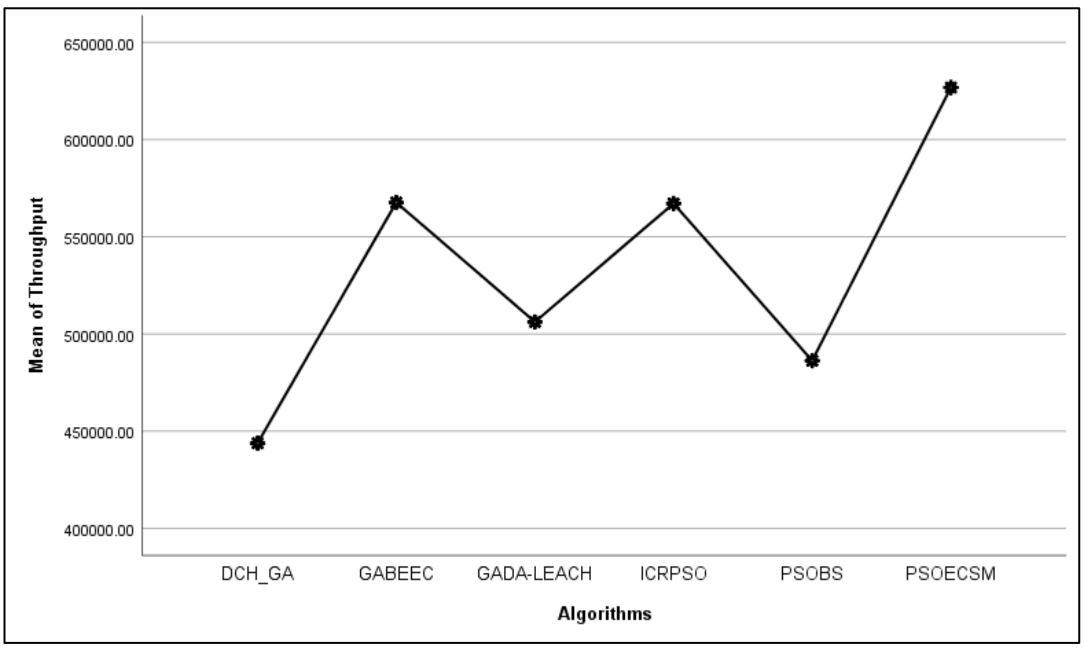


**Figure 12**. Estimated marginal mean plot for throughput with respect to each algorithm.

We conducted F-test is used to know whether two-samples may be regarded as drawn from the normal population have the same variance. The reason of applying F-test (based on analysis of variance (ANOVA)) is check that samples collected from six algorithms are within the same group or is there any possibility that one or the other algorithms shows different behaviour. Total 50 samples were drawn from each algorithm for both throughput and remaining energy values. The descriptive analysis is presented in Table 6 and 10 respectively for remaining energy and throughput and for each algorithm. Table 6 and 10 show that PSOECSM has achieved mean value (= 11.00) for remaining energy and mean value (= 626778.28) for throughput is higher than other algorithms. Table 7 and 11 presents the main ANOVA test results respectively for remaining energy and throughput. The significance value (p =

0.00 < 0.05) for both remaining energy and throughput. Hence, we could reject for both throughput and remaining energy. This result leads to a conclusion that one sample is better than the other one in the group. But which algorithm's sample is better is not indicated by ANOVA results. Therefore, we applied post hoc tests to compare the individual group by creating multiple pairs.

Results of two different post hoc tests (LSD and Tukey HSD) for both remaining energy and throughput is presented in Table 8 and 12. Both LSD and Tukey HSD tests are conducted because LSD test is very sensitive to the violation of the assumption of ANOVA. Hence, most likely to the Type 1 error (rejecting when it is true). To avoid this situation, Tukey HSD is performed which is less liable to the Type 1 error. The pairwise comparison between algorithms are presented in the Table 8 and 12 respectively for remaining energy and throughput. From this result we can conclude that the PSOECSM's performance is significantly better as compared to the other algorithms at 0.05 level of significance.

Homogeneity test is also conducted. Table 9 and 13 presents the result of homogeneity test for both remaining energy and throughput. Homogeneity test results reveals that the proposed PSOECSM is outperformed for throughput as well as remaining energy as it falls in a subset 6. This result also indicates that DCH-GA showed worst performance as far as throughput and remaining energy are concerned. Figure 11 and 12 graphically presented the mean value for remaining energy and throughput. The x-axis presents the different algorithms whilst y-axis shows the estimated marginal mean value. The graphical results also demonstrate the superiority of the proposed PSOECSM over other algorithms.

**4.4.1 Pitfalls of proposed work:** Although, the proposed work has shown the comprehensive improvement in the simulation scenario, we still feel there are a lot many challenges while implementing this algorithm in the real time scenarios.

Having not discussed the physical medium factors like fading, scattering and also the presence of obstacles, it will impose some challenges to the user. However, in the scope of this manuscript, the performance is optimized as far as routing of data packets is concerned.

The another pitfall of this approach is that it is not following secured wireless communication. Therefore, it is essential to impose security obligations to avoid any unauthenticated access to the network.

### 4.5 Summary

In nutshell, the summary of the improvement reported by PSO-ECSM is given in Table 14. The comparative analysis is done which indicates that PSO-ECSM outperform other algorithms in terms of different performance metrics.

**Table 14.** Comparative analysis of PSO-ECSM with other algorithms for different metrics.

| **Value of advanced fractions and quantity fractions of node** *$m=0.1, m_o=0.2, \beta=1, \alpha=2$* | | | | | |
|---|---|---|---|---|---|
| **Algorithms** | **Total Energy of Network (Joules)** | **Stability Period (rounds)** | **Half Node Dead (rounds)** | **Network Lifetime (rounds)** | **Throughput (packets)** |
| DCH-GA | 140 | 3608 | 7112 | 10466 | 308695 |
| GABEEC | 140 | 4390 | 9086 | 14840 | 409610 |
| PSOBS | 140 | 4370 | 8405 | 15222 | 486712 |
| GADA-LEACH | 140 | 4400 | 7722 | 12697 | 506203 |
| ICRPSO | 140 | 5048 | 9529 | 17360 | 568457 |
| **PSO-ECSM** | 140 | **6301** | **12558** | **19071** | **637880** |

**Table 15**. Comparative percentage improvement by PSO-ECSM to other algorithms.

| **Percentage (%) Improvement by PSO-ECSM algorithm** | | | | |
|---|---|---|---|---|
| Algorithms | Stability Period | Half Node Dead | Network Lifetime | Throughput |
| DCH-GA | 74.6 | 76.5 | 82.2 | 106.6 |
| GABEEC | 43.5 | 38.21 | 28.5 | 55.7 |
| PSOBS | 43.5 | 49.41 | 25.2 | 31 |
| GADA-LEACH | 43.2 | 62.62 | 50.2 | 26 |
| ICRPSO | 24.8 | 31.7 | 9.8 | 12.2 |

The percentage improvement by PSO-ECSM in terms of stability period, HND, Network Lifetime and Throughput is given in Table 15.

### 4.6 Result Analysis

It is observed from the simulation results that the proposed protocol PSO_ECSM reports the tremendous improvement in the different performance metrics namely, stability period, half network dead, network lifetime and throughput that too with the significant margin. After thorough inspection of simulation analysis, the specific reasons for such improvement is attributed to following reasons; the stability period is enhanced due to the optimized CH selection that is done through the five selection factors which helps in the preserving the energy of the sensor nodes. Furthermore, the optimized movement of the sink helps in the saving energy of CHs experiencing low energy whilst communication. The network lifetime is enhanced due to the combination of optimized CH selection as well as the optimized routing implemented through the PSO. The reception of data packets is also improved with a significant margin due to the strategic data collection performed by the sink. Therefore, the throughput in the proposed protocol is improved at a significant margin.

## 5. Conclusions

In this paper, we proposed PSO-ECSM method for both cluster head selection and sink mobility problem in the domain of wireless sensor network which can meet sustainability challenges in networks stability period, network lifetime, throughput etc., of real time issues. A multi-objective Particle Swarm Optimization (PSO) were employed in integration of fitness function in terms of different five parameters viz., residual energy, distance, node degree, average energy and Energy Consumption Rate (ECR). Furthermore, we have considered various related well-defined sink mobility factors like discovering lowest energy CH, distance to the CH, and size of cluster for optimizing of multi hop hot-spot problem using PSO. Finally, these considered parameters have been fed to PSO-ECSM algorithm and their performance have been evaluated on different existing algorithm alongside different performance metrics. From the experimental result, it is observed that the proposed PSO-ECSM achieve the higher performance compare to their counter parts. Moreover, the performance comparison with the existing competent schemes demonstrate the efficacy of the proposed scheme. The results of statistical test revealed that the proposed PSO-ECSM showed significantly better results as compared to other five competitive algorithms (ICRPSO, GADA-LEACH, PSOBS, GABEEC and DCH-GA). This evaluated can be used as a supportive mechanism for network system to validate their algorithm. In future scope, we will deliberate the alternative to the sink mobility will be sought to make the network more cost effective and whilst decreasing the end to end delay by addressing the hybrid approach.

## References


[1] H.M. Ammari, N. Gomes, W.I. Grosky, M. Jacques, B. Maxim, D. Yoon, Review of applications of wireless sensor networks, Wirel. Sens. Netw. Curr. Status Future Trends. (2012) 1.

[2] J. Henderson, C. Engel, S.-H. Minefields, Military Applications for Wireless Sensor Networks, (n.d.).

[3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Comput. Netw. 38 (2002) 393–422.

[4] M.E. Bayrakdar, Exploiting cognitive wireless nodes for priority-based data communication in terrestrial sensor networks, ETRI J. (2019).

[5] M.E. Bayrakdar, Energy-Efficient Technique for Monitoring of Agricultural Areas with Terrestrial Wireless Sensor Networks, J. Circuits Syst. Comput. (2019) 2050141.

[6] M.E. Bayrakdar, Cooperative communication based access technique for sensor networks, Int. J. Electron. 107 (2020) 212–225.

[7] S. Tanwar, N. Kumar, J.J. Rodrigues, A systematic review on heterogeneous routing protocols for wireless sensor network, J. Netw. Comput. Appl. 53 (2015) 39–56.

[8] Bhatia, T., Kansal, S., Goel, S., & Verma, A. K., A genetic algorithm-based distance-aware routing protocol for wireless sensor networks. *Computers & Electrical Engineering, 56 (2016)* 441–455.

[9] S. Verma, N. Sood, A.K. Sharma, Design of a novel routing architecture for harsh environment monitoring in heterogeneous WSN, IET Wirel. Sens. Syst. 8 (2018) 284–294.

[10] B. Singh, D.K. Lobiyal, A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks, Hum.-Centric Comput. Inf. Sci. 2 (2012) 13.

[11] T.-T. Nguyen, J.-S. Pan, T.-K. Dao, S.-C. Chu, Load balancing for mitigating hotspot problem in wireless sensor network based on enhanced diversity pollen, J. Inf. Telecommun. (2017) 1–16.

[12] B. Suh, S. Berber, Rendezvous points and routing path-selection strategies for wireless sensor networks with mobile sink, Electron. Lett. 52 (2015) 167–169.

[13] M.A.M. Vieira, C.N. Coelho, D.C. Da Silva, J.M. da Mata, Survey on wireless sensor network devices, in: EFTA 2003 2003 IEEE Conf. Emerg. Technol. Fact. Autom. Proc. Cat No 03TH8696, IEEE, 2003: pp. 537–544.

[14] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless microsensor networks, IEEE Trans. Wirel. Commun. 1 (2002) 660–670.

[15] S. Tyagi, N. Kumar, A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks, J. Netw. Comput. Appl. 36 (2013) 623–645.

[16] Elhoseny, M., Farouk, A., Zhou, N., Wang, M.-M., Abdalla, S., & Batle, J. (2017). Dynamic multi-hop clustering in a wireless sensor network: Performance improvement. *Wireless Personal Communications*, *95*(4), 3733–3753.

[17] G. Smaragdakis, I. Matta, A. Bestavros, SEP: A stable election protocol for clustered heterogeneous wireless sensor networks, Boston University Computer Science Department, 2004. http://open.bu.edu/handle/2144/1548 (accessed September 11, 2017).

[18] L. Qing, Q. Zhu, M. Wang, Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks, Comput. Commun. 29 (2006) 2230–2237.

[19] B. Elbhiri, R. Saadane, D. Aboutajdine, others, Developed Distributed Energy-Efficient Clustering (DDEEC) for heterogeneous wireless sensor networks, in: IV Commun. Mob. Netw. ISVC 2010 5th Int. Symp. On, IEEE, 2010: pp. 1–4. http://ieeexplore.ieee.org/abstract/document/5656252/ (accessed September 11, 2017).

[20] D. Kumar, T.C. Aseri, R.B. Patel, EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks, Comput. Commun. 32 (2009) 662–667.

[21] N. Javaid, T.N. Qureshi, A.H. Khan, A. Iqbal, E. Akhtar, M. Ishfaq, EDDEEC: Enhanced developed distributed energy-efficient clustering for heterogeneous wireless sensor networks, Procedia Comput. Sci. 19 (2013) 914–919.

[22] T.N. Qureshi, N. Javaid, A.H. Khan, A. Iqbal, E. Akhtar, M. Ishfaq, BEENISH: Balanced energy efficient network integrated super heterogeneous protocol for wireless sensor networks, Procedia Comput. Sci. 19 (2013) 920–925.

[23] M. Akbar, N. Javaid, M. Imran, N. Amjad, M.I. Khan, M. Guizani, Sink mobility aware energy-efficient network integrated super heterogeneous protocol for WSNs, EURASIP J. Wirel. Commun. Netw. 2016 (2016) 66.

[24] P.G.V. Naranjo, M. Shojafar, H. Mostafaei, Z. Pooranian, E. Baccarelli, P-SEP: A prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks, J. Supercomput. 73 (2017) 733–755.

[25] N. Mittal, U. Singh, Distance-based residual energy-efficient stable election protocol for WSNs, Arab. J. Sci. Eng. 40 (2015) 1637–1646.

[26] N. Mittal, U. Singh, B.S. Sohi, A stable energy efficient clustering protocol for wireless sensor networks, Wirel. Netw. 23 (2017) 1809–1821.

[27] N. Mittal, U. Singh, B.S. Sohi, A novel energy efficient stable clustering approach for wireless sensor networks, Wirel. Pers. Commun. 95 (2017) 2947–2971.

[28] S. Verma, N. Sood, A.K. Sharma, A novelistic approach for energy efficient routing using single and multiple data sinks in heterogeneous wireless sensor network, Peer--Peer Netw. Appl. 12 (2019) 1110–1136.

[29] S. Verma, N. Sood, A.K. Sharma, Genetic Algorithm-based Optimized Cluster Head selection for single and multiple data sinks in Heterogeneous Wireless Sensor Network, Appl. Soft Comput. (2019) 105788.

[30] S. Verma, N. Sood, A.K. Sharma, QoS provisioning-based routing protocols using multiple data sink in IoT-based WSN, Mod. Phys. Lett. A. 34 (2019) 1950235.

[31] D. Kumar, T.C. Aseri, R.B. Patel, EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks, Comput. Commun. 32 (2009) 662–667.

[32] A. Kashaf, N. Javaid, Z.A. Khan, I.A. Khan, TSEP: Threshold-sensitive stable election protocol for WSNs, in: Front. Inf. Technol. FIT 2012 10th Int. Conf. On, IEEE, 2012: pp. 164–168.

[33] P.G.V. Naranjo, M. Shojafar, H. Mostafaei, Z. Pooranian, E. Baccarelli, P-SEP: a prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks, J. Supercomput. 73 (2017) 733–755.

[34] N.A.A. Latiff, N.M.A. Latiff, R.B. Ahmad, Prolonging lifetime of wireless sensor networks with mobile base station using particle swarm optimization, in: 2011 Fourth Int. Conf. Model. Simul. Appl. Optim., IEEE, 2011: pp. 1–6.

[35] Y. Hu, Y. Ding, K. Hao, L. Ren, H. Han, An immune orthogonal learning particle swarm optimisation algorithm for routing recovery of wireless sensor networks with mobile sink, Int. J. Syst. Sci. 45 (2014) 337–350.

[36] B. Singh, D.K. Lobiyal, A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks, Hum.-Centric Comput. Inf. Sci. 2 (2012) 13.

[37] P. Kuila, P.K. Jana, Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach, Eng. Appl. Artif. Intell. 33 (2014) 127–140.

[38] J. RejinaParvin, C. Vasanthanayaki, Particle swarm optimization-based clustering by preventing residual nodes in wireless sensor networks, IEEE Sens. J. 15 (2015) 4264–4274.

[39] Q. Ni, Q. Pan, H. Du, C. Cao, Y. Zhai, A novel cluster head selection algorithm based on fuzzy clustering and particle swarm optimization, IEEE/ACM Trans. Comput. Biol. Bioinform. 14 (2015) 76–84.

[40] Y.-F. Hu, Y.-S. Ding, L.-H. Ren, K.-R. Hao, H. Han, An endocrine cooperative particle swarm optimization algorithm for routing recovery problem of wireless sensor networks with multiple mobile sinks, Inf. Sci. 300 (2015) 100–113.

[41] P.S. Rao, P.K. Jana, H. Banka, A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks, Wirel. Netw. 23 (2017) 2005–2020.

[42] M. Azharuddin, P.K. Jana, PSO-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks, Soft Comput. 21 (2017) 6825–6839.
[43] D.R. Edla, M.C. Kongara, R. Cheruku, A PSO Based Routing with Novel Fitness Function for Improving Lifetime of WSNs, Wirel. Pers. Commun. 104 (2019) 73–89.
[44] S. Tabibi, A. Ghaffari, Energy-Efficient Routing Mechanism for Mobile Sink in Wireless Sensor Networks Using Particle Swarm Optimization Algorithm, Wirel. Pers. Commun. 104 (2019) 199–216.
[45] T. Shankar, S. Shanmugavel, A. Rajesh, Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks, Swarm Evol. Comput. 30 (2016) 1–10.
[46] B. Pitchaimanickam, G. Murugaboopathi, A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks, Neural Comput. Appl. (2019) 1–15.
[47] K. Vijayalakshmi, P. Anandan, A multi objective Tabu particle swarm optimization for effective cluster head selection in WSN, Clust. Comput. 22 (2019) 12275–12282.
[48] D. Chandirasekaran, T. Jayabarathi, Cat swarm algorithm in wireless sensor networks for optimized cluster head selection: a real time approach, Clust. Comput. 22 (2019) 11351–11361.
[49] J. John, P. Rodrigues, MOTCO: Multi-objective Taylor Crow Optimization Algorithm for Cluster Head Selection in Energy Aware Wireless Sensor Network, Mob. Netw. Appl. (2019) 1–17.
[50] T.A. Alghamdi, Energy efficient protocol in wireless sensor network: optimized cluster head selection model, Telecommun. Syst. (2020) 1–15.
[51] G.P. Gupta, B. Saha, Load balanced clustering scheme using hybrid metaheuristic technique for mobile sink based wireless sensor networks, (n.d.).
[52] M. Krishnan, S. Yun, Y.M. Jung, Dynamic clustering approach with ACO-based mobile sink for data collection in WSNs, Wirel. Netw. 25 (2019) 4859–4871.
[53] R. Vijayashree, C. Suresh Ghana Dhas, Energy efficient data collection with multiple mobile sink using artificial bee colony algorithm in large-scale WSN, Automatika. 60 (2019) 555–563.
[54] H. Zhang, Z. Li, W. Shu, J. Chou, Ant colony optimization algorithm based on mobile sink data collection in industrial wireless sensor networks, EURASIP J. Wirel. Commun. Netw. 2019 (2019) 1–10.
[55] H. Wang, K. Li, W. Pedrycz, An Elite Hybrid Metaheuristic Optimization Algorithm for Maximizing Wireless Sensor Networks Lifetime with a Sink node, IEEE Sens. J. (2020).
[56] N. Gharaei, K.A. Bakar, S.Z.M. Hashim, A.H. Pourasl, Inter-and intra-cluster movement of mobile sink algorithms for cluster-based networks to enhance the network lifetime, Ad Hoc Netw. 85 (2019) 60–70.
[57] Bayraklı, S., & Erdogan, S. Z., Genetic algorithm-based energy efficient clusters (gabeec) in wireless sensor networks. *Procedia Computer Science, 10 (2012)* 247–254.
[58] Pandey, H.M., Chaudhary, A., Mehrotra, D. & Kendall, G., Maintaining regularity and generalization in data using the minimum description length principle and genetic algorithm: Case of grammatical inference. Swarm and Evolutionary Computation, 31 (2016) 11-23.
[59] Pandey, Hari Mohan, Ankit Chaudhary, and Deepti Mehrotra. "A comparative review of approaches to prevent premature convergence in GA." Applied Soft Computing 24 (2014): 1047-1077.
[60] Pandey, Hari Mohan. "Parameters quantification of genetic algorithm." *Information Systems Design and Intelligent Applications*. Springer, New Delhi, 2016. 711-719.