Chapter 3

MS based data collection in homogeneous networks

3.1 Introduction

In the Internet of Things (IoT) based smart systems, intelligent sensor nodes are used to communicate with other devices in the networks. Intelligence can be introduced to sensor nodes through the application of soft computing approaches. Therefore, Intelligence based sensor nodes are capable of building intelligent WSNs for real-time data gathering. Intelligent WSNs are an arrangement of intelligent sensor devices that are connected wirelessly. Sensor nodes are small and low-cost devices that can accumulate meaningful information from the environment and communicate with other nodes in the network [55]. Recently intelligent WSNs are preferred over traditional networks due to their wide area of application [97]. The intelligent WSNs can be useful in places where the use of the conventional network is challenging, high-priced, or inapplicable [98] such as underwater, space, smart city, environmental monitoring, and military zone [99, 100, 101, 102, 103].

Sensor nodes have limited battery capacity, so the network lifetime of WSNs is determined by the energy consumption rate of deployed sensor nodes. Replacing or recharging batteries in sensor nodes after deployment is not always possible [6]. Collecting data from the environment and transmitting that data to other nodes or sink are two prime operations in which energy is consumed in sensor nodes [104]. Consumption of energy at the time of sensing is constant. However, a sensor node's transmission energy varies depending on the node's location, and network topology [105]. Therefore, the lifetime of a network is affected by the data collection scheme. In most of the approaches, the static sink is used where data is transmitted through multi-hop communication [106]. Sink nodes are particular nodes that collect all the sensed information from other nodes. Due to multi-hop data transmission, nodes stationed near the sink convey a vast amount of traffic load compared to sensor nodes far away from the sink. Therefore, the sensor node stationed near the sink has higher energy expenditure compared to the far-away nodes from the sink [107]. These nodes die early and partition the whole network into distinct isolated sections that lead to the premature death of the network [46].

Data collection using Mobile Sink (MS) can reduce the energy consumption rate of the sensor nodes and improve network performance [108]. The mobile sink is a device that moves in the network and collects data from sensor nodes [109]. In MS based data collection approaches, MS visits individual sensor nodes to collect data that leads to high data collection delay. To speed up the data gathering process, a set of Rendezvous Points (RPs) are selected. RP is sojourn points where MS halts to gather data from sensor nodes [110]. Since MS visits a limited number of points to collect data, the data-gathering delay is reduced. Sensor nodes are organized into clusters and each cluster is assigned one RP. Sensor nodes send their data to MS when it visits their cluster's RP. Nodes send their data either directly to MS or via multi-hop communication.

Many existing approaches use single-hop data transmission [111, 44, 48], but having only single-hop communication between RP and sensor nodes leads to a longer route for MS. On the other hand, multi-hop data transmission approaches [49, 112, 58] use a data aggregation mechanism at intermediate nodes, which increases energy consumption and leads to the premature death of the network. This work proposes a Multi-Objective Gray Wolf Optimization (MOGWO) based effective data collection scheme. In the proposed MOGWO algorithm, the leader controls swarming to get the optimum solution for the problem. For the exploration and exploitation phases of swarm intelligence algorithms, the performance of the MOGWO is superior in the context of the standard test functions. In the proposed algorithm transmission distance of nodes, intermediate hop counts between RP and nodes, and the average distance among nodes belonging to the same cluster are jointly used to create optimal clusters and select the optimal position for RP within the network. Nodes send their data to MS when it visits their cluster's RP. To balance the energy consumption, the role of RP is rotated among nodes of the cluster based on residual energy. The simulation results are compared with the state of art approaches such as MOPSO [57], DC-ACO [50], EAPC [58], and WRP [52] algorithm. The results show that the proposed approach outperforms existing approaches in various criteria. The following are the main contributions of this chapter.

- A novel MOGWO based algorithm is developed to create optimal clusters and find an optimal set of RPs that significantly prevents premature death of the network.

- An optimal energy-efficient clustering algorithm is proposed that significantly reduces energy consumption and improves the network lifetime. Intermediate hop counts, transmission distance, and average distance among nodes of the same cluster are used as objective for multi-objective fitness function to form optimal clusters between the deployed sensor nodes.

- A novel intelligent mobile sink based data routing scheme is proposed where Pareto dominance is used to find non dominated solutions that significantly improve network lifetime and reduces transmission delay.

- A novel RP rotation scheme is proposed to balance the energy consumption between the deployed sensor nodes.

3.2 Network Model and Energy Model

3.2.1 Network model

In the proposed approach, $\mathcal{N}$ number of sensor nodes are randomly deployed in a A X A meter square region, and an MS is used for data collection from the deployed sensor nodes. MS moves with constant speed and has enough storage and energy capacity. Sensor nodes are homogeneous and static. Nodes have the same amount of initial energy, and a node is counted as dead when it drains its energy completely. The transmission distance between two nodes is the shortest distance after avoiding obstacles. The links between nodes are symmetric. RPs are selected in the network with the help of the proposed algorithm, and then MS visits each RP to collect data from sensor nodes and delivers it to BS. BS is aware of selected positions of RP and provides route information to the MS. Table 3.1 shows the terminologies used in this chapter. Fig. 3.1 presents the network model of the proposed approach.

Table 3.1: Terminologies and definitions

Terminologies	***Definition***
$\mathcal{N}$	Number of deployed sensor nodes.
$\mathcal{S}$	Set of sensor nodes. $\mathcal{S} = sn_1,\ sn_2,\ sn_3,\ sn_4.....sn_{\mathcal{N}}$.
r	Communication range of sensor nodes.
m	Number of rendezvous points.
$\mathcal{R}$	Set of rendezvous points. $\mathcal{R} = \mathcal{RP}_1,\ \mathcal{RP}_1,\ \mathcal{RP}_1....\mathcal{RP}_m$.
$rpid$	Id assigned to each RP. $rpid$ = 1,2,3,4....m. For eg. $rpid$ of $\mathcal{RP}_7$ is 7.
$\mathcal{T}Dist(sn_i)$	Euclidean distance between the sensor node and the next-hop node. Next-hop can be either a sensor node or an RP.
$\mathcal{D}ist(sn_i, sn_j)$	Euclidean distance between node sn_i and sn_j.
$\mathcal{H}Counts(sn_i)$	Number of hops required by sn_i to send data to RP.
$\mathcal{S}um\mathcal{D}(\mathcal{RP}_j)$	Sum of distances among nodes allocated to $\mathcal{RP}_j$
$allocate(sn_i)$	A variable that holds the $rpid$ of the RP to which the sensor node is allocated.
$\mathcal{AR}$	Archive which contains non-dominated solutions.

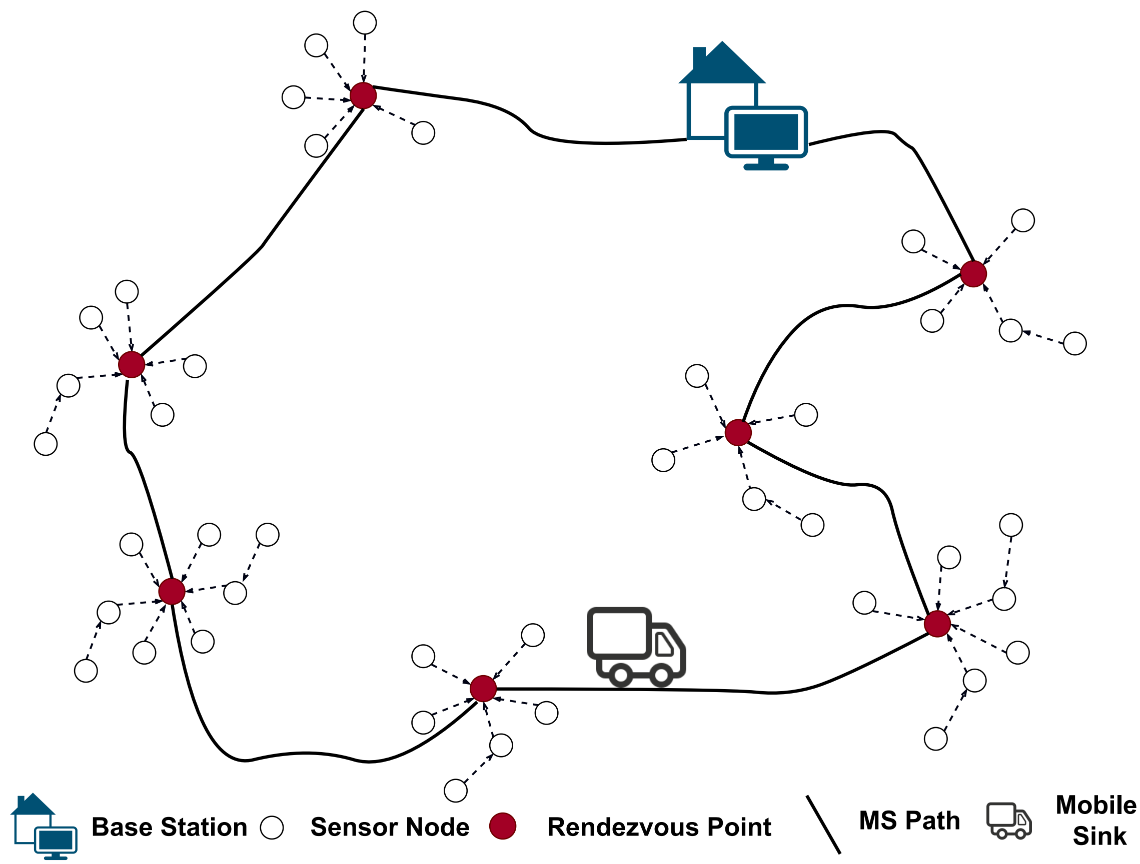


Figure 3.1: Rendezvous points based mobile sink traversal.

3.2.2 Energy model

In this work, the first order radio model is used for measuring the energy consumption of deployed sensor node. The energy model is same as [45]. Energy required for transmission of l bits of data over distance d is $\mathcal{E}_{tx}(l, d)$ and for receiving data is $\mathcal{E}_{rx}(l)$. The energy required for transmission can be divided into two parts, energy consumed in transmission circuitry and second is the energy consumed in the amplifier component.

$$\mathcal{E}_{(tx)}(l, d) = \begin{cases} l \times \mathcal{E}_{elec} + l \times \varepsilon_{fs} \times d^2, & d < d_o \\ l \times \mathcal{E}_{elec} + l \times \varepsilon_{mp} \times d^4, & d \geq d_o \end{cases} \tag{3.1}$$

$$\mathcal{E}_{rx}(l) = \mathcal{E}_{\text{elec}} \times l \tag{3.2}$$

where, $\mathcal{E}_{elec}$ is the energy required for electronic circuitry, ε_{fs} is the amplifier energy for free space and ε_{mp} is the amplifier energy for multipath fading channel. Distance d_o is the threshold distance, which is calculated as follows.

$$d_o = \frac{\varepsilon_{fs}}{\varepsilon_{mp}}$$

3.3 Multi-Objective Optimization and Pareto Optimality

A multi-objective problem holds two or more objectives. To get an optimal solution, it is required to satisfy several constraints [113]. A multi-objective minimization problem can be defined as

$$\begin{aligned} &\text{Minimize: } \mathcal{F}(\chi) = \{\mathcal{F}_1(x), \mathcal{F}_2(x), \ldots, \mathcal{F}_o(x)\}\,|x \in \mathcal{X} \\ &\text{Subject to: } \{\mathcal{C}_j(x) \leq 0, \quad j = 1, 2, \ldots, l\} \end{aligned} \tag{3.3}$$

where, o is the number of objectives, l is the number of constraints and $\mathcal{X}$ is a set of feasible solutions. In single-objective problems such as minimization problems, solution $\mathcal{A}$ is better than $\mathcal{B}$ only if $\mathcal{A} < \mathcal{B}$, so the solution can be found easily. In a multi-objective problem, there is more than one objective, so relational operators can not be applied directly. Hence in this work, Pareto optimality is used [114]. Here are a few definitions to understand the concept of Pareto optimality.

Pareto dominance: Let $\vec{a}$ and $\vec{b}$ are two vectors, then $\vec{a}$ is dominates $\vec{b}$ if and only if $\vec{a}$ is better than $\vec{b}$ for at least one objective and better or equivalent to $\vec{b}$ for rest of the objectives.

$$\forall i, \mathcal{F}_i(a) \leq \mathcal{F}_i(b) \wedge \exists i : \mathcal{F}_i(a) < \mathcal{F}_i(b) | 1 \leq i \leq o \tag{3.4}$$

Pareto optimal: A vector $\vec{a}$ is declared *Pareto optimal*, if and only if no other vector dominates $\vec{a}$.

Pareto optimal set: Pareto optimal set is set of non dominated feasible Pareto optimal solutions, which is denoted by $\mathscr{P}$.

$$\mathscr{P} = \{a \in \mathcal{X} | \nexists b \in \mathcal{X} : b \preceq a\} \tag{3.5}$$

Pareto front: The Pareto Front ($\mathscr{PF}$) is a set that contains the objective value of each solution in the Pareto optimal set.

$$\mathscr{PF} = \{\mathcal{F}(a) | a \in \mathscr{P}\} \tag{3.6}$$

3.4 Proposed Methodology

The proposed methodology is divided into two phases: the set-up phase and the intelligent data gathering phase. In the first phase, MOGWO based algorithm selects optimal RP position and groups nodes into optimal clusters. In the intelligent data gathering phase, an optimal data gathering path is designed to connect RPs. MS visits each RP using the optimal data gathering path and collects data from the deployed sensor nodes. At the end of each round node having the least residual energy is assigned RP's role in the cluster. Since every node's distance to its assigned RP is distinct, RP rotation balances the energy consumption, as the RP and nodes close to RP will consume less energy than the rest of the nodes. Fig. 3.2 shows the proposed intelligent data routing scheme.

3.4.1 Set-up phase

Sensor nodes send their data to MS either directly or in a multi-hop fashion. The selection of RP and clustering is done by performing MOGWO based algorithm. The objective is to minimize the maximum average transmission distance ($\mathscr{MaxATD}$), maximum average intermediate hop counts ($\mathscr{MaxAHcount}$) (between nodes and RP) and maximum average distance among nodes of the same cluster ($\mathscr{MaxACD}$) in the network.

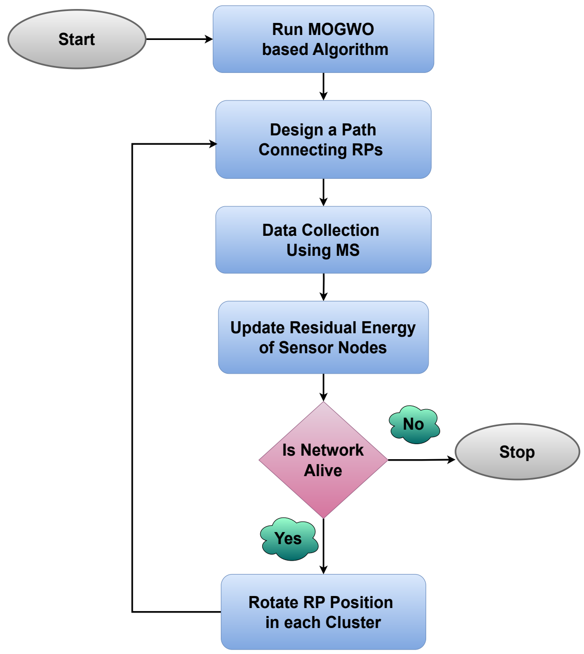


Figure 3.2: Proposed intelligent data routing scheme.

$\mathcal{MaxAHcount}$, $\mathcal{MaxATD}$ and $\mathcal{MaxACD}$ are calculated as follows. The total number of sensor nodes allocated to an RP, $\mathcal{RP}_j$

$$\mathcal{SumSN}(\mathcal{RP}_j) = \sum_{i=1}^{\mathcal{N}} t_i, \text{ where } t_i = \begin{cases} 1, & \text{if } \mathcal{allocate}(sn_i) = j \\ 0, & \text{otherwise} \end{cases} \tag{3.7}$$

The sum of transmission distances of sensor nodes allocated to $\mathcal{RP}_j$

$$\mathcal{SumTDist}(\mathcal{RP}_j) = \sum_{i=1}^{\mathcal{N}} d_i, \text{ where } d_i = \begin{cases} \mathcal{Tdist}(sn_i), & \text{if } \mathcal{allocate}(sn_i) = j \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

The average of transmission distance of nodes allocated to $\mathcal{RP}_j$

$$\mathcal{ATD}(\mathcal{RP}_j) = \frac{\mathcal{SumTDist}(\mathcal{RP}_j)}{\mathcal{SumSN}(\mathcal{RP}_j)} \tag{3.9}$$

In the network, the Maximum Average Transmission Distance $\mathcal{MaxATD}$ amidst the RPs is

$$\mathcal{MaxATD} = max\{\mathcal{ATD}(\mathcal{RP}_j)\}, 1 \leq j \leq m \tag{3.10}$$

The total sum of hop counts of the nodes which are allocated to $\mathcal{RP}_j$

$$\mathcal{SumHCount}(\mathcal{RP}_j) = \sum_{i=1}^{\mathcal{N}} h_i, \text{ where } h_i = \begin{cases} \mathcal{HCount}(sn_i), \text{ if } \mathit{allocate}(sn_i) = j \\ 0, \text{ otherwise} \end{cases} \tag{3.11}$$

The average hop count of sensor nodes allocated to $\mathcal{RP}_j$

$$\mathcal{AHCount}(\mathcal{RP}_j) = \frac{\mathcal{SumHCount}(\mathcal{RP}_j)}{\mathcal{SumSN}(\mathcal{RP}_j)} \tag{3.12}$$

The maximum average hop counts amidst the RPs is,

$$\mathcal{MaxAHcount} = max\{\mathcal{AHcount}(\mathcal{RP}_j)\}, 1 \leq j \leq m \tag{3.13}$$

The average distance among nodes allocated to $\mathcal{RP}_j$

$$\mathcal{ACD}(\mathcal{RP}_j) = \frac{\mathcal{SumD}(\mathcal{RP}_j)}{\mathcal{SumSN}(\mathcal{RP}_j)} \tag{3.14}$$

The maximum value of $\mathcal{SumD}(\mathcal{RP}_j)$ among RPs,

$$\mathcal{MaxACD} = max\{\mathcal{ACD}(\mathcal{RP}_j)\}, 1 \leq j \leq m \tag{3.15}$$

Once RPs are selected, sensor nodes are allocated to selected RPs and clusters are formed. Each sensor node is allocated to exactly one RP and each RP covers at least one sensor node. RP selection and clustering is done in such a fashion so that most of the sensor nodes present within one hop distance of an RP within the cluster. The sensor nodes that do not have an RP in single hop distance, are allocated to a cluster, with help of their closest sensor node.

Optimal cluster formation and RP selection: This approach proposes rendezvous point based path design for the mobile sink in such a way so that it minimizes $\mathcal{MaxATD}$, $\mathcal{MaxAHCount}$ and $\mathcal{MaxACD}$ at any rendezvous point. This problem of rendezvous point based path design for the mobile sink can be formulated as shown below.

$$\text{Minimize: Fitness } = \{\mathcal{MaxATD}, \mathcal{MaxAHCount}, \mathcal{MaxACD}\} \tag{3.16}$$

Such that:

$$\mathcal{Dist}(\mathcal{sn}_i, \mathcal{RP}_j) \leq r, \quad \forall \mathcal{sn}_i \in \mathcal{S} \text{ and } \forall \mathcal{RP}_j \in \mathcal{R} \tag{3.17}$$

$$\mathcal{Dist}(sn_i, sn_k) \leq r, \forall \mathcal{sn}_i, \mathcal{sn}_k \in \mathcal{S}, \forall \mathcal{RP}_j \in \mathcal{R} \text{ and } i \neq k \tag{3.18}$$

Equation 5.26 ensures that every sensor node is allocated to RP only if RP is present in the transmission range of the sensor node. Equation 5.27 ensures that if no RP is present in the transmission range of a sensor node, then that sensor node is allocated to an RP with the help of the nearest sensor node. For example, sn_1 does not have any RP in its communication range, but it has sn_2 which is allocated to $\mathcal{RP}_1$, then sn_1 will also be allocated to $\mathcal{RP}_1$ with the help of sn_2.

MOGWO: MOGWO is inherited from Gray Wolf Optimization (GWO) algorithm. GWO is based on the hunting and leadership behaviors of the gray wolves. The social hierarchy used by wolves is used as a mathematical model. Wolves have four categories. Alpha wolf, the fittest solution is considered as alpha, the beta wolf, which is the second-best solution, and third is delta wolf, third-best solution. These three wolves or solutions are called leader wolf and the rest of the solutions are called omega wolves. To imitate the circling behavior of wolves following equations are used [113].

$$\vec{\mathcal{D}} = \mid \vec{\mathcal{C}}.\vec{\mathcal{X}}_p(t) - \vec{\mathcal{X}}(t) \mid \tag{3.19}$$

$$\vec{\mathcal{X}}(t+1) = \vec{\mathcal{X}}_p(t) - \vec{\mathcal{A}}.\vec{\mathcal{D}} \tag{3.20}$$

In the above equations, t shows the current iteration, $\vec{\mathcal{A}}$ and $\vec{\mathcal{C}}$ are coefficients, $\vec{\mathcal{X}}_p$ represents the position vector of prey (in our case RP) and $\vec{\mathcal{X}}$ is the position vector of a gray wolf.

$$\vec{\mathcal{A}} = 2.\vec{a}.\vec{r}_1 - \vec{a}, \;\; \vec{\mathcal{C}} = 2.\vec{r}_2 \tag{3.21}$$

where $\vec{a}$'s value linearly decreases from 2 to 0 over the course of iterations and $\vec{r}_1$, $\vec{r}_2$ are random vectors in [0,1]. Value of $\vec{\mathcal{A}}$, between[-1,1] represents search agent converge toward prey. The value of coefficient $\vec{\mathcal{C}}$ varies between 0 to 2, which represents the random behavior of prey.

The algorithm simulates the behavior of wolves to find an optimal solution. The algorithm keeps the three fittest solutions and uses them to update other solutions. The following are the equations that are constantly run for each search agent.

$$\vec{\mathcal{D}}_\alpha = \mid \vec{\mathcal{C}}_1.\vec{\mathcal{X}}_\alpha - \vec{\mathcal{X}} \mid \tag{3.22}$$

$$\vec{\mathcal{D}}_\beta = \mid \vec{\mathcal{C}}_2.\vec{\mathcal{X}}_\beta - \vec{\mathcal{X}} \mid \tag{3.23}$$

$$\vec{\mathcal{D}}_\delta = \mid \vec{\mathcal{C}}_3.\vec{\mathcal{X}}_\delta - \vec{\mathcal{X}} \mid \tag{3.24}$$

$$\vec{\mathcal{X}}_1 = \vec{\mathcal{X}}_\alpha - \vec{\mathcal{A}}_1.\vec{\mathcal{D}}_\alpha \tag{3.25}$$

$$\vec{\mathcal{X}}_2 = \vec{\mathcal{X}}_\beta - \vec{\mathcal{A}}_2.\vec{\mathcal{D}}_\beta \tag{3.26}$$

$$\vec{\mathcal{X}}_3 = \vec{\mathcal{X}}_\delta - \vec{\mathcal{A}}_3.\vec{\mathcal{D}}_\delta \tag{3.27}$$

$$\vec{\mathcal{X}}(t+1) = \frac{\vec{\mathcal{X}}_1 + \vec{\mathcal{X}}_2 + \vec{\mathcal{X}}_3}{3} \tag{3.28}$$

Initialization of Gray Wolves: Each search agent holds a set of RP positions in the network. Each search agent has a fixed number of elements. These elements represent the RP set. The number of elements in a search agent is equal to the number of RPs. Initially, each element of a search agent x_i, i.e. $x_{i,j}$ (where $1 \leq i \leq p, 1 \leq j \leq m$) is assigned a uniform random number between 0 and 1 (here p is the size of population and m is the number of RP). In Search agent ($x_{i,j}$), value of j^{th} element represents position of a sensor node (sn_k) which is selected as RP. The following function is used to find the value of k.

$$k = Ceil(x_{i,j} \times n) \tag{3.29}$$

Archive management: MOGWO maintains an archive to store the non dominated solutions. A search agent x_i of $\mathcal{X}$ is inserted to the archive $\mathcal{AR}$, if and only if $\forall a_i \epsilon \mathcal{AR} \not\preceq x_i$. In words, a search agent x_i is added to $\mathcal{AR}$, if x_i is not dominated by any existing member of $\mathcal{AR}$. After inserting x_i to $\mathcal{AR}$, all existing members of $\mathcal{AR}$ that are dominated by x_i are removed from $\mathcal{AR}$.

Fitness Function: This study proposes the use of a multi-objective fitness function to find the best solution. There are three objectives. The first objective is to minimize the maximum average transmission distance of sensor nodes.

- Objective 1: Minimize $\mathcal{MaxATD}$

Second objective is to minimize the maximum average hop counts.

- Objective 2: Minimize $\mathcal{MaxAHCounts}$

The third objective is to minimize the maximum average transmission distance among sensor nodes of the same cluster.

- Objective 3: Minimize $\mathcal{MaxACD}$

These objectives are contrary in nature. When we try to decrease hop counts, the transmission distance between the sensor node and next hop node increases, and when we decrease the distance, the number of hops between RP and sensor node increases.

Algorithm 1: Proposed MOGWO based Algorithm

Output: Archive $\mathcal{AR}$ containing set of RPs

Initialize gray wolf population $x_i \epsilon \mathcal{X}$, $\forall i \leq 1 \leq m$
Initialize archive $\mathcal{AR} = \phi$
Initialize the variable $\vec{a}$, and coefficients $\vec{\mathcal{A}}$ and $\vec{\mathcal{C}}$
For each search agent, calculate $\mathcal{ATD}(\mathcal{RP}_j)$, $\mathcal{ACD}(\mathcal{RP}_j)$ and $\mathcal{AHcounts}(\mathcal{RP}_j)$ for each element.
Apply fitness function to determine the objective values ($\mathcal{MaxATD}$, $\mathcal{MaxACD}$ and $\mathcal{MaxAHcount}$) of each search agent
Discover the non-dominated solutions and store them in the $\mathcal{AR}$
$\mathcal{X}_a = Select\_Leader(archive)$ and remove α from $\mathcal{AR}$ temporarily
$\mathcal{X}_b = Select\_Leader(archive)$ and remove β from $\mathcal{AR}$ temporarily
$\mathcal{X}_d = Select\_Leader(archive)$ and and put α and β back to $\mathcal{AR}$
for *j=1 to maxIteration* **do**
    Update the position of each search agent by applying equations 3.22 to 3.28
    Update the values of a, $\vec{\mathcal{A}}$, and $\vec{\mathcal{C}}$
    For each search agent, calculate $\mathcal{ATD}(\mathcal{RP}_j)$, $\mathcal{ACD}(\mathcal{RP}_j)$ and $\mathcal{AHcounts}(\mathcal{RP}_j)$ for each element.
    Determine the objective values of each search agent
    Discover the non-dominated solutions
    Update the $\mathcal{AR}$ concerning the obtained non-dominated solutions
    $\mathcal{X}_a = Select\_Leader(archive)$ and remove α from $\mathcal{AR}$ temporarily
    $\mathcal{X}_b = Select\_Leader(archive)$ and remove β from $\mathcal{AR}$ temporarily
    $\mathcal{X}_d = Select\_Leader(archive)$ and put α and β back to $\mathcal{AR}$
return the archive $\mathcal{AR}$
Pick a suitable solution from $\mathcal{AR}$

Hence, we can write a fitness function as

$$Fitness = Minimize\{\mathcal{MaxATD}, \mathcal{MaxAHcount}, \mathcal{MaxACD}\} \tag{3.30}$$

The proposed algorithm selects RP and creates cluster with optimal values for $\mathcal{MaxATD}$, $\mathcal{MaxAHcount}$, and $\mathcal{MaxACD}$.

3.4.2 Proposed MOGWO based algorithm

The proposed MOGWO-based algorithm is shown in algorithm 1. The first step of the algorithm is the initialization, where the grey wolf population, the archive $\mathcal{AR}$, and variables $\vec{a}$, $\vec{\mathcal{A}}$, $\vec{\mathcal{C}}$ are initialized. These grey wolves are the search agents, and each grey wolf contains an array of potential positions for RP in the network. Next, the fitness function is used to calculate the objective value of each search agent. The objective values are nothing but $\mathcal{MaxAHCounts}$, $\mathcal{MaxATD}$ and $\mathcal{MaxACD}$. Based on the obtained objective values, the Pareto dominance is applied to search agents to find non-dominated search agents. These search agents are placed in the archive. α, β, and δ wolves are selected from the agents present in the archive. These three agents are the fittest solutions present in the archive.

From equation 3.22 to 3.28 are used to update position of each search agent. Next values of variables $\vec{a}$, $\vec{\mathcal{A}}$, $\vec{\mathcal{C}}$ are updated by using equation 3.21. Objective values for all search agents are calculated, and archive $\mathcal{AR}$ is updated with non dominated solutions again. From the archive, α, β, and δ wolves are selected. The above steps are repeated until the maximum iteration is reached. When the maximum iteration is reached, archive $\mathcal{AR}$ is returned, and a suitable solution is selected from the archive. All the solution present in $\mathcal{AR}$ are optimal solutions. The solution contains the set of positions that are selected as RP.

3.4.3 Intelligent routing phase

After the selection of RPs, a path or trajectory needed to be designed for effective data routing. The path is the sequence of RPs, an order in which RPs are visited by MS. To form a path TSP is applied. When MS reaches an RP, it broadcasts a *Hello* message along with $\mathcal{rpid}$ within cluster. The sensor nodes that receive *Hello* message from the MS check the received $\mathcal{rpid}$. If they are allocated to the same RP, then they transmit the Acknowledgement ($\mathcal{ACK}$) message to the MS. After receiving $\mathcal{ACK}$ messages from all cluster nodes, MS allocates a data transmission time slot to each sensor node. According to the given time slot each sensor node transmits its data to the MS by minimum energy consumption. Residual energy of each node is updated after every rounds. Based on residual energy role of RP is rotated among cluster nodes.

Table 3.2: Simulation parameters

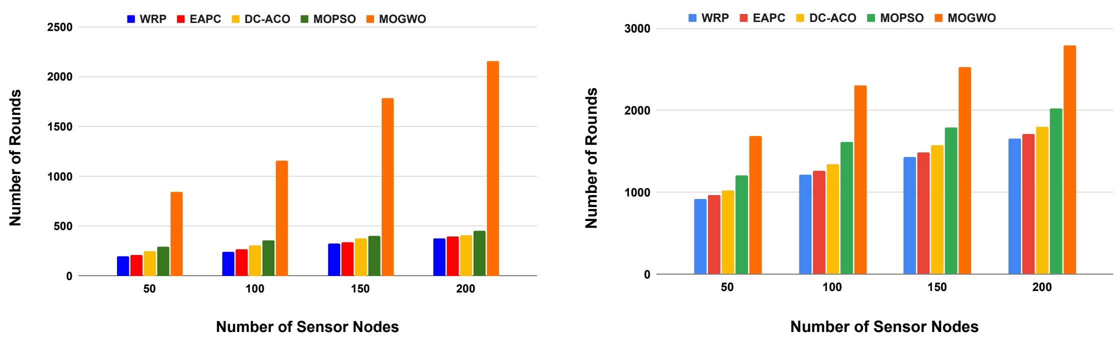
Parameters	***Values***
Target area	100 X 100 m^2
Number of sensor nodes	50-200
Initial energy of each sensor node	0.5J
$\mathcal{E}_{\text{elec}}$	50nJ/bit
ε_{fs}	10pJ/bit
ε_{mp}	0.0013pJ/bit
Gray Wolves population size	30
Maximum Number of iterations	100
$\vec{a}$	2 to 0
$\vec{r}_1$, $\vec{r}_2$	[0,1]

3.5 Performance Analysis

This section analyses the performance of the proposed MOGWO based algorithm through various simulations. We have conducted simulations on ns-3. Table 3.2 shows the parameters used in the simulation. We have considered a deployment area of 100 X 100 square meters for WSNs simulation. Uniform random distribution is used for the deployment of nodes. The mobile sink is assumed to have adequate power and storage. The simulation results are compared with MOPSO [57], DC-ACO [50], EAPC [58], and WRP [52] algorithms.

3.5.1 Stability period and network lifetime

The stability period refers to the time when first node dies in the network. A node is considered dead when it drains its energy completely. In the WSNs, with limited battery for sensor nodes higher stability period is major requirement. Fig. 3.3a depicts the stability period of the network with different node densities. Since the target area is fixed, with varying node topology, inter-distance among cluster nodes, transmission distance of sensor nodes, and hop counts also change. The proposed approach minimizes the distance and hops between RP and sensor nodes, leading to less energy consumption in the network, which also changes the network stability. The proposed scheme improves networks stability 80% than WRP, 78% than EAPC, 75% than DC-ACO, and 72% than MOPSO. The proposed scheme creates optimal clusters with optimal RP positions.



(a) Stability period. (b) Network lifetime.

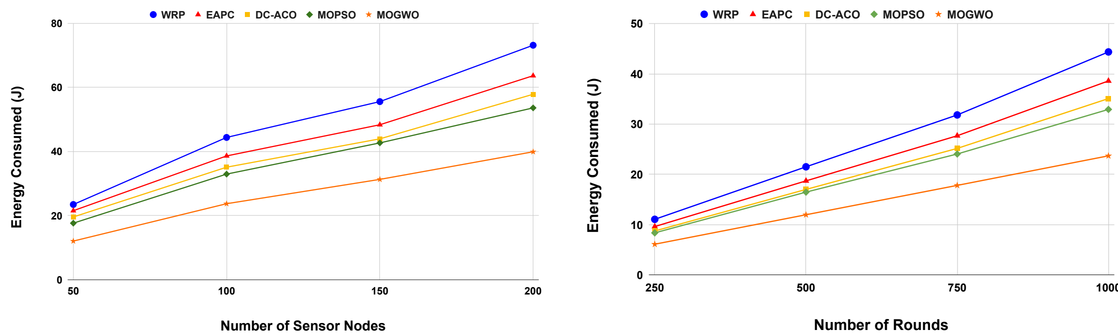
Figure 3.3: Stability period and network lifetime.

With minimized inter-node distance of the cluster, the RP rotation scheme balances the energy consumption. This leads to longer stability period in the network. This also indicates that the proposed scheme effectively prevents premature death of the network.

Fig. 3.3b shows the network lifetime in terms of the number of rounds with different node densities. In this study, we consider a network dead when a certain percentage of nodes die in the network. Fig. 3.3b shows that the proposed MOGWO improves network lifetime by 71%, 67%, 55% and 38% compared to WRP, EAPC, DC-ACO and MOPSO respectively on average. Network lifetime is directly affected by the lifetime of sensor nodes. Thus reducing energy consumption at sensor nodes will directly help in increasing the overall lifetime of the network.

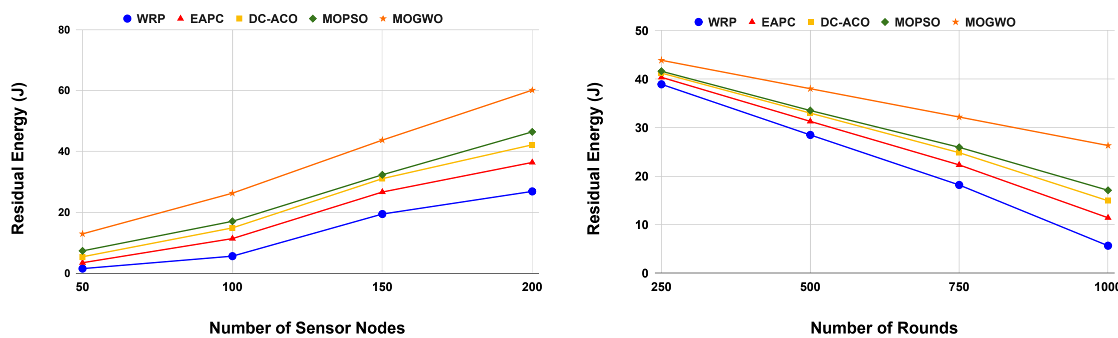
3.5.2 Energy consumption and residual energy

This experiment evaluates the amount of energy consumed in the network. Fig. 3.4a shows total energy consumed for varying node densities after 1000 rounds and Fig. 3.4b shows total energy consumed with respect to number of rounds. These results shows that WRP consumes 76%, EAPC consumes 61%, DC-ACO consumes 46% and MOPSO consumes 37% more energy than Proposed scheme. In WSNs, data transmission is prime cause of energy consumption. The proposed algorithm creates optimal clusters with optimal RP positions, which minimizes the hop counts and transmission



(a) With varying node densities. (b) With respect to rounds.

Figure 3.4: Energy consumption.



(a) With varying node densities. (b) With respect to rounds.

Figure 3.5: Residual energy.

distance. Hence nodes need less energy for data transmission. Also, rotation of RP within cluster leads more balanced energy consumption. Fig. 3.5a and Fig. 3.5b show residual energy of the network with respect to varying node densities and number of rounds respectively. These figures indicate that proposed MOGWO based approach saves more energy at sensor nodes than WRP, EAPC, DC-ACO and MOPSO approach. The proposed approach minimizes transmission distance so less energy is consumed in data transmission. This approach reduces intermediate hops which minimizes data aggregation at intermediate hops. Furthermore, RP rotation causes uniform energy consumption in clusters which leads to higher residual energy in the network.

3.5.3 Alive nodes

This experiment evaluates the number of alive nodes present in the network relating to number of rounds. Node is considered alive when it has sufficient energy to perform all its operations. Fig. 3.6 shows the number of alive nodes at different rounds. Number of alive nodes decreases with increase in number of rounds. The simulation results indicate that the proposed scheme has the highest number of alive nodes. This improvement is seen due to the proposed MS-based intelligent data gathering scheme where the optimal clusters are created, and optimal RP positions are selected for energy-efficient data gathering. The nodes consume less energy, which keeps them alive for more rounds.

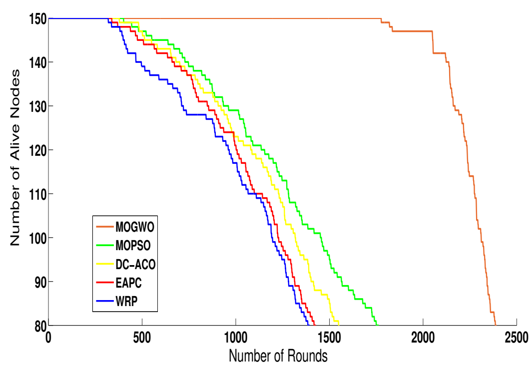


Figure 3.6: Number of alive nodes.

3.6 Summary

In this chapter, an MS-based data collection scheme for homogeneous WSNs has been proposed. The proposed scheme considers hop counts, transmission distance, and distance among deployed sensor nodes for cluster formation and an optimal number of RP selections. To achieve this a Multi-Objective Gray Wolf Optimization mechanism is used. Furthermore, this work also applies an RP rotation technique to balance the energy consumption among the cluster member nodes. It helps to prevent early death of the network. Extensive simulations have been conducted to demonstrate the performance of the proposed scheme. Simulation results show that the proposed scheme outperforms as compared to the state-of-the-art approaches in terms of stability period, network lifetime, energy consumption, residual energy, and number of alive nodes in the network. The results indicate that the performance of the proposed scheme is superior as compared to the state-of-the-art algorithms.

Advanced IoT applications use heterogeneous sensor nodes for data collection from physical objects. Heterogeneous sensor nodes are different in sensing, communication, and battery capacity. These heterogeneous sensor nodes create heterogeneous WSNs for data gathering. Heterogeneous networks are more complex in nature as compared to homogeneous networks. Data routing algorithms that are designed for heterogeneous WSNs must consider variations in communication, sensing and battery capacity during the protocol design and development. Therefore, the next chapter presents a deep policy dynamic programming-based intelligent data routing scheme for heterogeneous WSNs.