



# An energy optimized and QoS concerned data gathering protocol for wireless sensor network using variable dimensional PSO

Saugata Roy<sup>a,\*</sup>, Nabajyoti Mazumdar<sup>b</sup>, Rajendra Pamula<sup>a</sup>

<sup>a</sup> Department of Computer Science & Engineering, Indian Institute of Technology (ISM) Dhanbad, 826004, India

<sup>b</sup> Department of Information Technology, Indian Institute of Information Technology Allahabad, 211012, India

## ARTICLE INFO

### Keywords:

Ant colony optimization  
Collection point  
Mobile data collector  
Quality of service  
Variable dimensional particle swarm optimization  
Wireless sensor network

## ABSTRACT

This paper presents a quality of service (QoS) concerned energy efficient data gathering protocol for wireless sensor network (WSN) by means of a mobile data collector (MDC). Several resource constraints of WSN nodes make the QoS preservation very challenging, particularly the energy constraint. Moreover, the energy hole problem during data communication has emerged as a major threat to QoS preservation in WSN. The state-of-the-art algorithms have employed MDC as a promising approach to effectively alleviate the energy hole problem. Efficient trajectory design can allow the MDC to sojourn at limited data collection points (CPs) while covering significantly more nodes, reducing energy consumption and latency in data collection. However, estimation of the optimal number of CPs and MDC trajectory design is an NP-hard problem in a large-scale network. To solve this problem, variable dimensional particle swarm optimization (VD-PSO) is leveraged for the optimal nomination of CPs, and ant colony optimization is applied to determine the best trajectory visiting all the nominated CPs. A multi-objective fitness function is designed to assess the quality of the nominated CPs by considering various QoS metrics. In order to accelerate the convergence of the VD-PSO algorithm, a modified particle update procedure and dimensionality pruning are also introduced. The simulation results depict the effectiveness of the proposed protocol in terms of network lifetime, total energy consumption, end-to-end delay, packet delivery ratio, etc.

## 1. Introduction

In recent years, Internet of Things (IoT) has emerged as a paramount approach to couple smart devices via network and perform data exchanges among them [1,2]. One of the significant technology of IoT is Wireless Sensor Network (WSN) [3,4], a collection of small embedded systems referred to as wireless sensor nodes having sensing, processing, communication, and power units. These nodes are usually deployed over a region of interest to measure various physical parameters and periodically report them to the remote sink or Base Station (BS) for further processing. Optimization of energy consumption has evolved as a crucial objective in WSN, as nodes are placed in hard-to-reach environments resulting in their batteries non-rechargeable [5,6]. In addition to the energy preservation of the sensor nodes, satisfying the network Quality of Service (QoS) requirements like end-to-end reliability and data delivery timeliness are other major concerns of several mission-critical and real-time multimedia applications [7–10]. Under the circumstances, the notion of clustering [11–16] has been extensively embraced by the majority of the researchers to improve the network lifetime and data transfer reliability. In a clustered WSN, a representative of each cluster called Cluster Head (CH) is selected, which

is responsible for cluster data collection, aggregation, propagation, etc. However, in a static sink scenario, cluster based communication protocols inflict a huge relay load on the CHs in the vicinity of the sink, which may lead to their rapid energy exhaustion and premature death. This is popularly known as the energy hole problem [17,18] that leads to a deterioration in the network QoS by diminishing the network lifespan, packet delivery ratio, and data delivery timeliness. The introduction of a Mobile Data Collector (MDC) [19–31] has evolved as a promising technique to successfully alleviate the energy hole problem. MDC is a movable unit in a clustered WSN responsible for collecting the cluster data from the CHs by moving around the deployment region and ultimately delivering the data collected to the BS.

Traditional MDC-based data gathering methods can be divided into two key groups on the basis of whether or not MDC visits each sensor node for data gathering. In the first group [20–22], MDC visits individual sensor nodes for data acquisition, resulting in better energy preservation and collision-free data transmission. Such a methodology allows the sensor nodes to perform data transmission obeying the free-space energy model [14] due to shorter transmission distance.

\* Corresponding author.

E-mail address: [saugata.2016dr57@cse.ism.ac.in](mailto:saugata.2016dr57@cse.ism.ac.in) (S. Roy).

Nevertheless, this technique may not be feasible in delay-sensitive applications of WSN as it produces a higher number of data Collection Points (CPs) that elongate the MDC tour. The other school of thought [23–29] enables the MDC to sojourn at a restricted number of CPs where every sensor node forwards its data to the nearest CP. Such a restricted data collection increases the average data transmission distance to the MDC, which enforces the sensor nodes to obey the multi-path energy model or multi-hop routing for data delivery. However, as stated by the first order radio model [14], the multi-path model incurs rapid energy dissipation with respect to the transmission distance. On the contrary, multi-hop routing may give rise to the energy hole probability around the CPs. Hence, this data acquisition scheme does not guarantee an energy-saving architecture despite reducing the data delivery latency significantly. Considering the foregoing incompatibility between energy conservation and data delivery latency, MDC based data gathering protocols ought to leverage the benefits of both energy and delay-sensitive approaches. This encourages the present-day researchers to explore an optimal set of MDC collection points that ensures free-space data transmission from the sensor nodes.

Nominating such an optimal set of CPs using an exhaustive search method causes an NP-hard problem in a large-scale WSN. Solutions to the NP-hard problems include searching for potential solutions across the large spaces. Population based methodologies of swarm intelligence [32–34] have been commonly used on a variety of these issues. Particle Swarm Optimization (**PSO**) is one of the widely explored swarm intelligence algorithms stimulated by the social behaviour of birds and fishes. In PSO, a swarm of particles builds a solution space where each particle is a potential solution to the multidimensional search space. Due to its effectiveness in solving NP-hard problems, PSO can be adopted to optimize the CP nomination for an MDC based WSN [35–38]. Herein, the particle dimensionality is decided by the number of CPs present in the MDC trajectory. Therefore, every particle in the swarm represents one potential MDC trajectory, i.e., a sequence of CPs. However, in practice, it is not possible to precisely predict the optimal number of CPs beforehand, which prevents the traditional PSO from producing an optimal solution. Considering this fact, the necessity of an improved variant of PSO called Variable Dimensional Particle Swarm Optimization (**VD-PSO**) has emerged in MDC based sensor data gathering. VD-PSO initiates its iterative search process with different dimensional particles (potential solution) and eventually converges to the optimal dimensional solution. Evaluating the optimal set of CPs is followed by the construction of least cost MDC trajectory to visit all the CPs exactly once and return back to the starting point. Such a minimal length hamiltonian cycle formulation from a fully connected weighted graph again experiences a non-polynomial time solution for large-scale WSNs. This motivates modern-day scholars to adopt a construction based optimization strategy like Ant Colony Optimization (**ACO**) that is mimicked by the foraging behaviour of the real ants. In WSN, ACO is generally applied to approximate the MDC trajectory from a set of predefined CPs in polynomial time [39–41].

### 1.1. Motivation

In an energy-constrained delay-sensitive WSN, MDC based data gathering brings the following key challenges:

- **Energy:** A plethora of research work [4–6] reveals that prolonging the network lifetime is the utmost goal in an energy-constrained network. Considering the energy consumption rate, the CHs prefer to communicate with their respective CPs following the free-space energy model. This necessitates the need for optimal CP nomination such that all proximate CHs around the CPs are covered by the free-space communication range.
- **Latency:** Limiting the communication range of the CHs may increase the number of CPs, which in turn elongates the MDC trajectory. Therefore, acquiring the free-space data transmission

model in MDC based data gathering is a challenging task while considering the delay-sensitive applications of WSN. On the contrary, data transmission obeying the free-space model maximizes the network lifetime by conserving the network energy (Eq. (1)). Hence, the protocols designed for MDC based routing should aim at designing an optimal trajectory to attain a trade-off between energy preservation and data delivery latency.

- **Optimal number of CPs:** Most of the literature on MDC based data gathering have aimed to nominate the optimal position of  $k$  number of CPs. However, it is inconvenient to predict the suitable value of  $k$  in a large-scale WSN since changes in the network topology may demand a variation in the  $k$  value. More concretely, inaccurate estimation of the  $k$  value may result in either of the following issues:

- ✓ Overestimation: Results in elongated MDC tour; accordingly, increases the data acquisition latency.
- ✓ Underestimation: Does not ensure energy efficiency as it produces either multipath or multi-hop data transmission.

In this context, the variable dimensional PSO is found to be more suitable to estimate the optimal number of CPs ( $k$ ).

### 1.2. Contribution

The aforementioned challenges encourage us to propose a QoS aware energy-efficient data gathering scheme by designing an optimal MDC trajectory with the help of multi-objective VD-PSO algorithm. Unlike traditional PSO, the proposed VD-PSO algorithm begins its iterative searching with a diversified population and ultimately converges to the optimal solution comprising an optimal number of CPs with their optimal locations. The multi-objective fitness function of the proposed VD-PSO algorithm leverages both energy and QoS sensitive metrics to assess the quality of each candidate solution. In a nutshell, the contribution of this article is given below :

- Discovering an optimal set of CPs by means of the VD-PSO algorithm.
- Constructing a minimal length MDC trajectory that passes through each CP exactly once.
- Presenting a novel particle updation procedure to update the velocity and position of the particles having different dimensionalities.
- Designing a multi-objective fitness function to efficiently address the incompatibility among energy aware QoS routing metrics.
- Performing an extensive simulation analysis under Gaussian and uniformly distributed WSN with different node densities to demonstrate the feasibility and dominance of the proposed protocol over the existing ones.

## 2. Related work

Over the past two decades, several research works [4–6] reveal that optimizing the energy consumption of the sensor nodes is a paramount concern in energy-constrained applications of WSN. Besides energy optimization, network Quality of Service (QoS) metrics have evolved as key requisites of any mission-critical, and real-time multimedia applications [7–10]. Cluster based communication protocols [11–16] have shown better competency in terms of energy-saving and reliability over the traditional routing protocols. However, in static sink scenario, cluster based multi-hop routing may encounter the energy hole issue [17,18]. In the circumstances, recent studies [19–31] recommend the notion of MDC for sensory data gathering as the most promising approach in order to successfully mitigate the energy hole problem. This literature survey divides MDC based data gathering schemes into two major classes:

## 2.1. Point-to-point-data-forwarding

This class of data gathering scheme assures an energy-saving architecture that allows the MDC to visit individual sensor nodes for data collection. Study [20] develops Mobile Sink (MS) assisted data collection strategy by means of a progressive optimization technique that takes the data rate constraints between sensor nodes and MS into account. The major objective of this strategy is minimizing the data collection delay by combining the proximate collection sites, actively omitting the redundant nodes, and finally identifying the appropriate start and finish positions of data collection. Authors in [21], solves the Delivery Latency Minimization Problem (DLMP) by developing the Substitution Heuristic Algorithm (SHA) that discovers the anchor points within the communication radius of the sensor nodes as well as eliminates the redundant anchor points. This guarantees a better shortening of the MS travel path. Study [22] presents a reduced TSP path for the MS, where it visits some nominated anchor points to gather the sensory data. These anchor points establish the MS travelling path passing through the communication radius of each sensor node. Such an improved TSP route results in better energy minimization for MS based data collection. Nevertheless, the above-mentioned studies suggest data gathering from each sensor node that lengthens the MDC travel path; consequently may not be feasible in delay-sensitive WSN applications.

## 2.2. Partial-data-forwarding

The other class of data gathering scheme secures lower data delivery latency by allowing the MDC to stop at a limited number of CPs for data gathering. Each sensor node forwards its data to the nearest CP via direct or multi-hop communication. Depending on the MDC trajectory type, the approaches that fall in this class can be further categorized as:

### 2.2.1. Predetermined MDC trajectory

The studies [23–26] fall in this category suggest that the MDC moves along a predetermined trajectory to collect the sensory data throughout the network operation. In [23], an MDC based Adaptive Immune Energy Efficient clustering Protocol (MSIEEP) is constructed to mitigate the energy hole problem. To minimize the hop counts and dropped packets, MSIEEP partitions the deployment area into a number of equal-sized regions, each of which leads to one MDC sojourn location. Furthermore, MSIEEP evaluates the optimal number of CHs along with their locations utilizing the Adaptive Immune Algorithm (AIA). Study [24] introduces a rendezvous point (RP) based data acquisition protocol that notably mitigates the energy hole problem. In this article [24], a virtual grid based infrastructure is designed to nominate the MS sojourn locations while the RPs are chosen from each grid cell utilizing the fuzzy interference system. Another distributed RP selection protocol is proposed in article [25] which takes both energy as well as coverage-aware parameters into account for the clustering and routing process. This sort of methodology ensures minimal hop data transmission between the sensor nodes and the respective MS sojourn points, thus minimizes the data collection latency. Authors in [26] geographically cellularize the whole deployment region and employ two mobile sinks responsible for collecting the cell agent data. This type of framework substantially minimizes the energy dissipation and data delivery delay of the network. However, despite reducing the MDC tour length significantly, all the studies in this category do not adapt to dynamic changes in the network topology.

### 2.2.2. Random MDC trajectory

In this category, the data collector trajectory is random in nature that is built based on the network topology. In [28], an Energy-Aware Path Construction (EAPC) algorithm is suggested. The principle is to consider the path cost between consecutive Data Collection Points

(DCPs) to obtain an improved path of data collection. The EAPC algorithm is initiated by the Minimum Spanning Tree (MST) construction, followed by the nomination of a suitable set of DCPs, and ultimately allows the MDC to accomplish data collection from heavily loaded DCPs. In spite of the formation of an improved data collection path, EAPC may experience a long-chain multi-hop routing under large-scale unbalanced deployments. A further study [29] on sink trajectory design allows the network to dynamically choose the RPs based on the distributed clustering technique. Authors in [30,31] propose two cellular infrastructure based routing protocols, namely Routing based on Cellular structure and Clustering (RCC) and Routing Based on Grid structure and Mobile sink (RBGM). RCC and RBGM provide efficient data dissemination from the cell members to the mobile sink with the least possible energy expenditure. Furthermore, they introduce an optimal routing strategy that updates the cell headers with the latest sink position instead of advertising it throughout the network, thus achieving a lower data dissemination delay.

However, the foregoing MDC based routing approaches induce a high computational complexity while designing an optimal MDC travelling path in large-scale WSN. This motivates present-day scholars to explore the bio-inspired meta-heuristics [32–34] intensely in order to optimize the MDC traversal path in polynomial time. Particle Swarm Optimization (PSO) is one of the frequently used metaheuristics exploited by several researchers [35–38] for WSN based routing problems. Utilizing the single objective PSO, study [35] introduces an effective RP selection strategy under two equality constraints, like data transmission latency and traffic rate constraints. Refs. [36,37] on the other side, exploit multi-objective PSO and design an optimal MDC path planning that aims to shorten the MDC trajectory for data collection. To obtain a better minimization in tour length, the proposed EETP [37] protocol guides the MDC to sojourn at potential visiting points within the overlapping of communication ranges instead of passing through the centre of the sensor nodes. More recently, a PSO-based Energy efficient Clustering and Sink Mobility (PSO-ECSM) algorithm is addressed in the study [38] that optimizes the CH nomination procedure by means of five energy sensitive factors.

Calculation of optimal CPs is followed by the construction of a minimal cost MDC trajectory that infers nothing but a minimal length hamiltonian cycle, which cannot be solved in polynomial time. Such an NP-hard problem can be competently addressed by Ant Colony Optimization (ACO) algorithm [39–41] which built a near-optimal solution in polynomial time. Authors in [39] develop a dynamic clustering approach that enhances the network lifetime by distributing the dissipated energy load among the sensor nodes. Subsequently, exploiting ACO algorithm it builds an optimal path for the MDC. An ACO based Mobile Sink Path Determination (ACO-MSPD) algorithm is presented in the article [40] that identifies a near-optimal RP set by leveraging the forwarding load of the sensors under non-uniform data generation.

In spite of providing several productive thoughts about MS trajectory design, there is a research gap in the existing approaches to address the incompatibility between energy preservation and data delivery delay. In this context, this article aims at attaining a well balance between energy preservation and delay by nominating an optimal number of CPs. The selection of such an optimal number of CPs is achieved by the novel VD-PSO method, and the trajectory of the MDC for visiting the CPs is established by the Ant Colony Optimization algorithm. In Table 1, we highlight various key characteristics of several existing protocols and the proposed one.

## 3. System model

### 3.1. The WSN model

This paper considers a WSN framework where a set of sensor nodes  $S_{sn} = \{s_1, s_2, \dots, s_n\}$  and an MDC are deployed over a target region of size  $\mathcal{A} = \mathcal{L} \times \mathcal{L} \text{ unit}^2$ . Here,  $\mathcal{L}$  denotes the length of the target region

**Table 1**  
Summary and comparison of related works.

Protocol	Method applied	Sink trajectory	# of CPs	Energy consumption	Delay	Hotspot probability
SAS [20]	TSP heuristic	Random	variable	Low	High	Minimal
SHA [21]	TSP heuristic	Random	Variable	Low	High	Minimal
MSIEEP [23]	Immune algorithm	Predetermined	Fixed	Medium	Medium	Low
EGRPM [26]	Distributed	Predetermined	Fixed	High	High	Medium
VGRSS [24]	Distributed	Predetermined	Fixed	Medium	High	Low
Hierarchical [25]	Distributed	Predetermined	Fixed	Medium	Medium	Low
RCC [30]	Distributed	Random	Variable	Medium	Medium	Low
RBGM [31]	Distributed	Random	Variable	Medium	Medium	Low
MOPSO [36]	PSO	Random	Fixed	Medium	High	Low
PSO-ECSM [38]	PSO	Random	Fixed	Medium	Medium	Medium
ACO based [39]	ACO	Random	Variable	Medium	High	Medium
ACO-MSPD [40]	ACO	Random	Variable	Medium	Low	Low
Proposed	VD-PSO, ACO	Random	Variable	Low	Low	Minimal

side. The proposed protocol initiates its working on a clustered WSN where the cluster head selection is achieved by following a distributed clustering algorithm proposed in [25]. In article [25], node's residual energy, neighbour centrality, and coverage significance parameters are considered for CH nomination where the set of selected CHs can be denoted as  $\mathcal{S}_{ch} = \{S_1, S_2, \dots, S_{n_{ch}}\}$  where  $\mathcal{S}_{ch} \subset \mathcal{S}_{sn}$ . Each CH accumulates its member data with its own sensory data in order to deliver them to the respective MDC. The MDC moves through the target region and stops to collect the CH data at specific CPs. Herein, a maximum Allowed Tour Time ( $ATT_{max}$ ) of the MDC is taken as data collection threshold to meet the delay-sensitive requirements of the WSN. Following are the basic assumptions employed throughout this article:

- All deployed sensor nodes remain static throughout their lifetime.
- All sensor nodes are aware of their geographical position using GPS (Global Positioning System) or any well-known localization algorithms [42,43].
- All sensors are identical in terms of battery power, sensing radius, and communication radius.
- The power level of each sensor node can be controlled in such a way that maximum power level can be used for direct data transmission to the MDC [44,45].
- All sensor nodes are time synchronized [46].
- The mobile data collector is a resource-rich device, i.e., it has no energy and processing constraints.

### 3.2. First order radio model

To estimate the energy consumption during data communication and aggregation among the sensor nodes this article utilizes the first order radio model as described in [14]. The energy expended by the radio at transmitter and receiver ends are pictorially presented in Fig. 1 where  $\epsilon_{amp}$  and  $E_{elec}$  are the energy dissipation per bit in the amplifier and electronic circuit, respectively. We symbolize the exponent of path loss by  $e$  whose value depends on the transmission distance  $d$ . If distance  $d$  is less than threshold distance  $d_{th}$  then according to free-space ( $fs$ ) model  $e$  is set to 2 and  $\epsilon_{amp} = \epsilon_{fs}$  otherwise, according to multipath ( $mp$ ) model  $e$  is set to 4 and  $\epsilon_{amp} = \epsilon_{mp}$ . The threshold distance is derived as  $d_{th} = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$ .

Hence, the transmission energy equation can be rewritten as

$$E_{Tx}(k, d) = \begin{cases} E_{elec} * k + \epsilon_{fs} * k * d^2 & \text{if } d < d_{th} \\ E_{elec} * k + \epsilon_{mp} * k * d^4 & \text{if } d \geq d_{th} \end{cases} \quad (1)$$

### 3.3. Applied optimization methods (VD-PSO and ACO)

In this subsection, we present a brief overview of the optimization methods employed in the proposed article. In order to nominate the optimal set of CPs a novel variant of particle swarm optimization called VD-PSO is adopted, whereas to form the trajectory of the MDC connecting all the CPs, ant colony optimization is incorporated. The mentioned optimization methods are discussed below:

#### 3.3.1. Variable dimensional PSO (VD-PSO)

Particle Swarm Optimization (PSO), designed by James Kennedy and Russell C. Eberhart (1995) is a bio-inspired metaheuristic optimization technique that addresses computationally hard problems in a reasonable amount of time. In basic PSO, each particle is a potential solution that searches for the optimum solution in the  $D$ -dimensional search space. The position and velocity of  $i$ th particle  $P_{(i,D)}$  is expressed as  $X_i = (X_{i1}, X_{i2}, \dots, X_{iD})$  and  $V_i = (V_{i1}, V_{i2}, \dots, V_{iD})$ , respectively. The best position of  $P_{(i,D)}$  found from its previous experience ( $pbest$ ) is represented as  $Pb_i = (Pb_{i1}, Pb_{i2}, \dots, Pb_{iD})$  and the best position found in the whole swarm ( $gbest$ ) is  $Gb = (Gb_1, Gb_2, \dots, Gb_D)$ . At every time step  $t$ , each particle continuously updates its velocity and position with respect to the  $pbest$  and  $gbest$  values as follows:

$$V_{id}^{t+1} = w * V_{id}^t + c_1 * rand(0, 1) * (Pb_{id}^t - X_{id}^t) + c_2 * rand(0, 1) * (Gb_d^t - X_{id}^t) \quad (2)$$

$$X_{id}^{t+1} = X_{id}^t + V_{id}^t \quad (3)$$

where  $w$  is the inertia weight that regulates the balance between swarm local and global exploration ability. Constants  $c_1, c_2$  are the cognitive weight (particle best influence) and social weight (swarm overall best influence) respectively. The iterative updation continues until the maximum number of iterations ( $t_{max}$ ) is reached or an adequate  $gbest$  value is obtained.

As one of the most commonly used optimization techniques in WSN, PSO usually helps to seek optimal points in the large search space (deployment area); accordingly, it is widely employed in the mobile sink based routing problem. The goal of this article is to establish an MDC trajectory that comprises an optimal set of CPs. Traditional PSO algorithm relates each of such possible paths to a  $D$ -dimensional particle wherein every dimension represents the location of one tentative CP on that path. However, with traditional PSO, it is impracticable to estimate the precise number of CPs as there is no diversion in particle length. Usually, PSO initiates its searching process with a predefined number of CPs that induces either overestimation or underestimation and consequently degrades the network QoS. This necessitates the Variable Dimensional Particle Swarm Optimization (VD-PSO) to be emerged to address the foregoing issue (as shown in Fig. 2). Since particles in the VD-PSO swarm are of variable dimensionality, the traditional updating formula for velocity and position (Eqs. (2) and (3)) is not feasible here. Hence, a novel particle updation strategy with variable dimensionality is proposed in this article (presented in Section 5).

#### 3.4. Ant colony optimization

Ant Colony Optimization (ACO), proposed by M. Dorigo et al. (1996), is a prevalent bio-inspired metaheuristic that is imitated by the food searching strategy of the natural ants. ACO uses a number

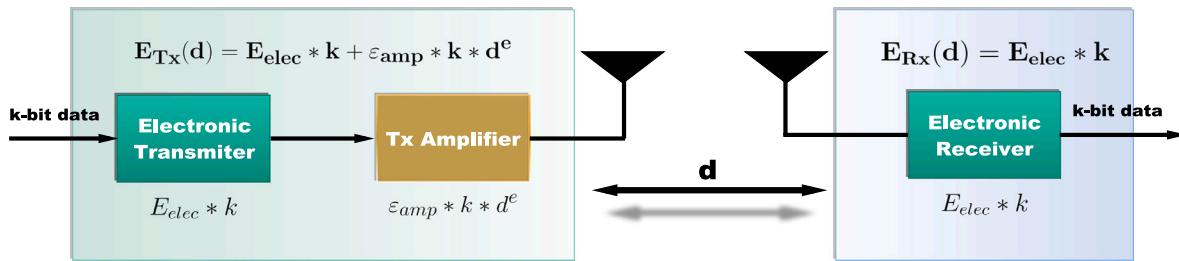


Fig. 1. First order radio model.



Fig. 2. Traditional PSO vs. VD-PSO.

of artificial ants with distinctive advantages over natural ants as (i) the artificial ants possess memory, i.e., they will not visit the CP that has been visited already, (ii) they are aware of the distance among the CPs and prefer to choose the nearest CP if the pheromone level is same, and (iii) if the lengths of two paths are the same, the path with more pheromones would appear to be chosen by the ants. ACO is a construction based strategy that aims at searching for the minimal length hamiltonian cycle in the solution space. The solution space of the ACO is a fully connected weighted graph. Over the past few years, it has been heavily exploited in the field of WSN to search for an optimal traversal cycle among a set of predefined nodes in polynomial time.

#### 4. Mathematical model

This section comprises the problem formulation and the design of the VD-PSO fitness function, which are vividly analyzed in the following subsections.

##### 4.1. Problem formulation

In this subsection, we present two formulations of the proposed methodology. The first formulation aims to optimize the number of CPs while ensuring connectivity to all CHs within free-space communication range. This results in a balance between energy consumption

and data delivery delay. The minimization problem of CPs can be formulated as:

$$\text{P1 : } \underset{n_{cp}}{\text{minimize}} \quad n_{cp} \quad (4a)$$

$$\text{subject to } \sum_{i=1}^{n_{cp}} B_{ij} \geq 1, \quad \forall j \in \mathcal{S}_{ch} \quad (4b)$$

where  $n_{cp}$  is the number of CPs in the WSN. Constraint (4b) states that for each CH  $j \in \mathcal{S}_{ch}$  there exists atleast one CP to obtain free-space connectivity.  $B_{ij}$  is a binary variable and can be defined by:

$$B_{ij} = \begin{cases} 1 & \text{dist}(i, j) \leq d_{th} \\ 0 & \text{otherwise} \end{cases}$$

where  $\text{dist}(i, j)$  is the euclidean distance between CH  $j$  and CP  $i$ ,  $d_{th}$  is the threshold distance for free-space communication range (as defined in Section 3.2). The second formulation of the proposed protocol involves in constructing an optimal MDC trajectory to visit every CP exactly once and returning back to the BS. This sort of optimal trajectory design mainly focuses on reducing the total tour distance of the MDC, and consequently mitigates the data gathering delay. This is none but a  $n_{cp}$  city travelling salesman problem and can be formulated

as:

$$\mathbf{P2} : \text{minimize} \quad \sum_{i=1}^{n_{cp}} \sum_{j=1, j \neq i}^{n_{cp}} C_{ij} B_{ij} \quad (5a)$$

$$\text{subject to} \quad \sum_{i=1, i \neq j}^{n_{cp}} B_{ij} = 1, \quad \forall j = 1, \dots, n_{cp} \quad (5b)$$

$$\sum_{j=1, j \neq i}^{n_{cp}} B_{ij} = 1, \quad \forall i = 1, \dots, n_{cp} \quad (5c)$$

$$\mathcal{V}_i - \mathcal{V}_j + n_{cp} B_{ij} \leq n_{cp} - 1 \quad 2 \leq i \neq j \leq n_{cp} \quad (5d)$$

$$\mathcal{V}_i \geq 0 \quad \forall i = 1, \dots, n_{cp} \quad (5e)$$

$$B_{ij} \in \{0, 1\} \quad \forall i, j = 1, \dots, n_{cp} \quad (5f)$$

where  $C_{ij}$  is the travelling distance to visit CP  $j$  from CP  $i$  and  $B_{ij}$  is a binary variable. Constraints (5b) and (5c) refers to exactly one entry for a CP  $j$  and exactly one exit for a CP  $i$ , respectively. Eq. (5d) refers to the subtour elimination constraint where  $\mathcal{V}_i$  is the auxiliary variable  $\forall i = 1, \dots, n_{cp}$ .

#### 4.2. Fitness function design

This paper aims to design an optimal MDC trajectory that will competently satisfy the QoS routing requirements in WSN. However, the following research questions have emerged as major challenges while designing such MDC trajectory

- How to determine the optimal number of CPs with their optimal locations?
- How to evaluate the shortest MDC path that passes through each nominated CP?

The above challenges are addressed herein with the exploitation of proposed VD-PSO and basic ACO algorithms, respectively (Section 5). The proposed VD-PSO algorithm introduces a multi-objective fitness function to evaluate the quality of each particle. Each sub-objective of the fitness function corresponds to one QoS metric, and their combined effect needs to be optimized such that the QoS routing requisites are achieved. To regulate the relative influence of the conflicting sub-objectives, the Weighted Sum Model (WSM) is applied during the construction of the fitness function (Eq. (13)). The fundamental properties of the following sub-objectives are explored as:

##### 4.2.1. Energy efficiency

Prolonging the network lifetime is the primary goal in an energy optimized WSN. Furthermore, this article is only concerned to know how the cluster data are forwarded to the MDC. Hence, the energy aware QoS routing for MDC based data gathering ought to be designed in such a way that the average transmission energy among the CHs and the MDC is minimized. The fitness of each particle in terms of energy efficiency can be represented by the Average Transmission Energy (ATE) which is calculated as follows:

$$ATE = \frac{\sum_{i=1}^{n_{ch}} E_{TX}(i, j)}{\sum_{i=1}^{n_{ch}} E_{res}(i)} \quad (6)$$

where  $ATE \in [0, 1]$ ,  $n_{ch}$  is the number of CHs,  $E_{TX}(i, j)$  is the energy consumption due to data transmission between CH  $S_i \in S_{ch}$  and its respective CP  $j$ ,  $E_{res}(i)$  is the residual energy of CH  $S_i$ .

##### 4.2.2. Packet delivery

Successful delivery of data packets to the MDC implies the end-to-end reliability of the network and consequently improves the QoS routing requirements. In real-world applications, the number of received packets is always lesser than the number of transmitted packets because of heavy congestion, error-prone medium, hostile environment, and so on. Therefore, the fitness function aims to maximize the average

Packet Delivery Ratio (PDR) value of the network and designs its second sub-objective as given below:

$$PDR = \frac{\sum \text{No. of packets received by MDC}}{\sum \text{No. of packets transmitted}} \quad (7)$$

##### 4.2.3. End-to-end delay

The third metric for QoS routing is measured in terms of end-to-end delay ( $D_{e2e}$ ) of the network which can be expressed as a cumulative sum of the link delays. Link delay  $D_L$  is the time (in ms) required to successfully deliver a packet from the sender node to the receiver node of the communication link.  $D_L$  involves transmission delay ( $D_{tran}$ ), propagation delay ( $D_{prop}$ ), processing delay ( $D_{proc}$ ), and queuing delay ( $D_{que}$ ). In other words, link delay between sensor nodes  $s_i$  and  $s_j$  can be mathematically written as:

$$D_L(i, j) = D_{tran}(i, j) + D_{prop}(i, j) \times \text{packet\_size} + D_{proc}(i, j) + D_{que}(i, j) \quad (8)$$

Here,  $D_{tran}(i, j)$  is the time needed to send out all the bits of a packet from the sensor node  $s_i$  into the communication link and can be defined as:

$$D_{tran}(i, j) = \frac{n_{bits}^i}{R_{tran}}$$

where  $n_{bits}$  is the packet size (in bits) generated at sensor node  $s_i$  and  $R_{tran}$  is the data transmission rate (bits/sec) of the outgoing link.  $D_{prop}$  is the time needed for a packet bit to travel from the sender node  $s_i$  to the receiver node  $s_j$  of the communication link. It can be defined as:

$$D_{prop}(i, j) = \frac{Dist(i, j)}{v_{prop}}$$

where  $Dist(i, j)$  is the distance (metres) between  $s_i$  and  $s_j$  and  $v_{prop}$  is the velocity of the propagation (metre/sec).  $D_{que}(i, j)$  is the wait time of a packet in the receiver queue before being processed by the controller of the receiving node  $s_j$ . The value of  $D_{que}$  depends on three factors namely (i) size of the receiver queue, (ii) incoming packet arrival rate ( $R_{arr}$ ), and (iii) data transmission rate of the outgoing link ( $R_{tran}$ ). Hence, the traffic intensity can be measured as the ratio of  $R_{arr}$  and  $R_{tran}$ , i.e.,  $\tau_{intn} = \frac{R_{arr} \cdot n_{bits}^i}{R_{tran}}$  where  $n_{bits}^i$  is the packet size at  $s_i$ . It is obvious that if  $\tau_{intn} > 1$  then  $D_{que}$  exists otherwise, it is 0.  $D_{proc}(i, j)$  refers to the time required to process the packet header of a packet by the controller of the receiving node  $s_j$ . The value of  $D_{proc}$  depends on the processing speed of the controller. Herein, it is assumed to be negligible and thus not considered in link delay computation. Hence, Eq. (8) can be rewritten as:

$$D_L(i, j) = D_{tran}(i, j) + D_{prop}(i, j) \times n_{bits}^i + D_{que}(i, j) \quad (9)$$

Now, from the aforesaid definition and Eq. (9) we can mathematically express the average end-to-end delay of the network as:

$$D_{e2e} = \sum_{i=1}^n \sum_{x \in \{j, I_i, CP_i\}} D_L(i, x) \quad (10)$$

where  $n$  denotes the number of sensor nodes,  $j$  denotes the immediate next hop of  $s_i$ ,  $I_i$  denotes the set of intermediate nodes between  $s_i$  and respective CP  $CP_i$ . It is noteworthy that a higher value of  $D_{e2e}$  may cause discarding of the packets, which further results in packet retransmission and higher energy consumption. This suggests the fitness function of VD-PSO to minimize the average end-to-end delay (as defined in Eq. (10)) and design it in a normalized way as:

$$Delay = \frac{D_{e2e}}{ATT_{max}} \quad (11)$$

where  $Delay \in [0, 1]$  and  $ATT_{max}$  is the maximum allowed tour time of the MDC.

#### 4.2.4. Link quality

One of the crucial requisites of QoS routing is network link quality that is usually indicated by the hardware metric Received Signal Strength Indicator (**RSSI**). RSSI is a measure of how well the receiver end can hear an RF signal from the sender. According to the RF transceiver CC2420 typical RSSI value varies in  $[-100, 0]$  dBm where larger value denotes stronger signal received. In this study, the proposed fitness function aims to optimize the link quality among CHs and the MDC, where the average link quality of the network can be calculated as:

$$LQ_{avg} = \frac{\sum_{i=1}^{n_{ch}} RSSI(i, j)}{n_{ch} \times \min RSSI} \quad (12)$$

where  $LQ_{avg} \in [0, 1]$ ,  $n_{ch}$  is the number of CHs, and  $RSSI(i, j)$  is the received signal strength indicator between sender CH  $S_i \in S_{ch}$  and its respective CP  $j$ . The lower the  $LQ_{avg}$  value, the better the link quality is.

The aforementioned sub-objective functions are conflicting in nature as the fitness function will be optimized only when the first, third, and fourth objective functions are minimized and the second one is maximized. Therefore, all four objectives are combined with the help of WSM to build the final fitness function (**Fit**) that ought to be minimized.

$$Fit = W_1 \times ATE + W_2 \times (1 - PDR) + W_3 \times Delay + W_4 \times LQ_{avg} \quad (13)$$

where  $W_1$ ,  $W_2$ ,  $W_3$ , and  $W_4$  are the weight factors to regulate the contribution of each sub-objective function.

## 5. Proposed work

Considering the challenges discussed in the previous section, we propose an optimal MDC trajectory design protocol in QoS routing environment. The working of the proposed protocol is divided into three phases, namely, (i) CP nomination, (ii) MDC trajectory construction, and (iii) Steady phase (as shown in Fig. 3). Determining the optimal number of CPs with their optimal locations is accomplished by exploiting the VD-PSO algorithm. During the CP nomination phase, the following tasks are accomplished:

- Initializing the particles with variable lengths to begin a search of different dimensionalities.
- Evaluation of particle quality by means of a multi-objective fitness function.
- Presenting an improved particle updation procedure to update their velocities and positions.
- Pruning the irrelevant dimensions of higher dimensional particles in order to accelerate the converge of the VD-PSO search.

Next, in the MDC trajectory construction phase, the shortest possible MDC travelling cycle is established to visit each nominated CP exactly once. Such an optimal trajectory is obtained by leveraging the ACO algorithm that approximates the best solution in polynomial time. Finally, the steady phase is responsible for gathering the network data at the MDC collection points. The working of each phase is vividly described in the following subsections:

### 5.1. CP nomination

This phase presents a stepwise description of the proposed VD-PSO operation that produces an optimal set of CPs.

#### 5.1.1. Initialization of particles

Since the optimal number of CPs may vary with respect to the deployment scenario, VD-PSO swarm comprises particles with different dimensionalities. Each particle dimensionality is equal to the number of CPs on the potential MDC path described by that particle. In other words, if there are  $k$  number of particles in the swarm, there ought to be  $k$  possible paths. The  $j$ th path  $P_{(j,n_j)}$  can be represented as a sequence of ordered pairs as  $P_{(j,n_j)} = \langle (x_{j1}, y_{j1}), (x_{j2}, y_{j2}), \dots, (x_{jn_j}, y_{jn_j}) \rangle$  where each ordered pair is a CP location coordinates and  $n_j$  is the number of the CPs in the  $j$ th path. Correspondingly, the  $j$ th particle  $P_{(j,n_j)}$  can be represented as:

$$\boxed{(x_{j1}, y_{j1})} \quad \boxed{(x_{j2}, y_{j2})} \quad \dots \dots \quad \boxed{(x_{jn_j}, y_{jn_j})}$$

where each component of the particle is an ordered pair of coordinates, and the particle dimensionality ( $Dim_j$ ) is  $n_j$ . However, assigning different dimensionality to each particle in the swarm may diminish the efficiency of VD-PSO algorithm as the particles cannot learn much from each other when they are extremely different. Hence, the proposed protocol divides the whole swarm into a number of groups of equal dimensionality particles. Each group has  $\frac{N_s}{N_g}$  number of particles where  $N_s$  is the swarm size and  $N_g$  is the number of groups. This is referred to as **Swarm Division** and presented as:

$$\underbrace{P_1, \dots, P_{\frac{N_s}{N_g}}, P_{\frac{N_s}{N_g}+1}, \dots}_{1^{\text{st}} \text{ group}}, \underbrace{P_{2 \cdot \frac{N_s}{N_g}}, \dots, \dots, P_{N_s - \frac{N_s}{N_g} + 1}, \dots, P_{N_s}}_{2^{\text{nd}} \text{ group}}, \dots, \underbrace{\dots, \dots, P_{N_s - \frac{N_s}{N_g} + 1}, \dots, P_{N_s}}_{N_g^{\text{th}} \text{ group}}$$

Now, each component of a particle is randomly initialized by means of assigning a uniformly distributed random number in the range  $[1, \mathcal{L} - 1]$  to each  $x$  and  $y$  coordinate where  $\mathcal{L}$  is the length of the target region side.

#### 5.1.2. Evaluation of particles

The quality of each particle is evaluated with the help of a multi-objective fitness function (Eq. (13)) that assigns a fitness score to each particle. The fitness score of a particle defines how well it addresses the QoS routing constraints for the constructed MDC trajectory, where a lower fitness score indicates a better candidate solution. The fitness function helps each particle to iteratively find its personal best ( $pbest$ ) and the swarm global best ( $gbest$ ). It should be noted that if two particles having the same fitness score made a tie for the  $gbest$ , then the one with lower dimensionality will be chosen as  $gbest$ .

#### 5.1.3. Updation of particles

Unlike traditional PSO, VD-PSO deals with variable dimensional particles, thus proposes a novel particle updation method. This method suggests that each component of a particle is updated by finding the nearest position (component) in its  $pbest$  and the swarm  $gbest$ . Suppose that the  $k$ th component of the particle  $P_{(j,n_j)}$  needs to be updated in the current iteration  $t$ . So, every component of the respective  $pbest$  is scanned to find the nearest one to the  $k$ th component of  $P_{(j,n_j)}$ . Likewise, all the components of  $gbest$  are also checked. Suppose, for  $pbest_j$ , component  $l$  is found to be the nearest one to the  $k$ th component of  $P_{(j,n_j)}$  and for  $gbest$  it is component  $m$ . Now, particle  $P_{(j,n_j)}$  can update its velocity and position using Eqs. (14) and (15)

$$\begin{aligned} V_{jk}^{t+1} &= w * V_{jk}^t + c_1 * rand(0, 1) * (Pb_{jl}^t - X_{jk}^t) \\ &\quad + c_2 * rand(0, 1) * (Gb_m^t - X_{jk}^t) \end{aligned} \quad (14)$$

$$X_{jk}^{t+1} = X_{jk}^t + V_{jk}^t \quad (15)$$

It should be noted that after updating the particle positions, some updated locations may lie beyond the range  $[1, \mathcal{L} - 1]$ , i.e., the  $x$  and  $y$

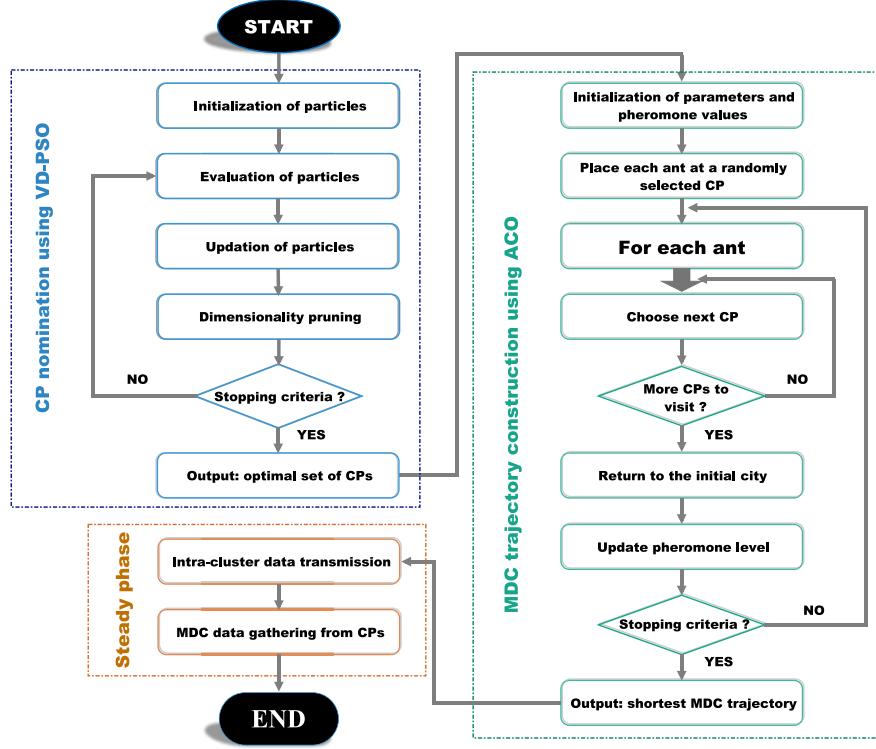


Fig. 3. Flow diagram of the proposed protocol.

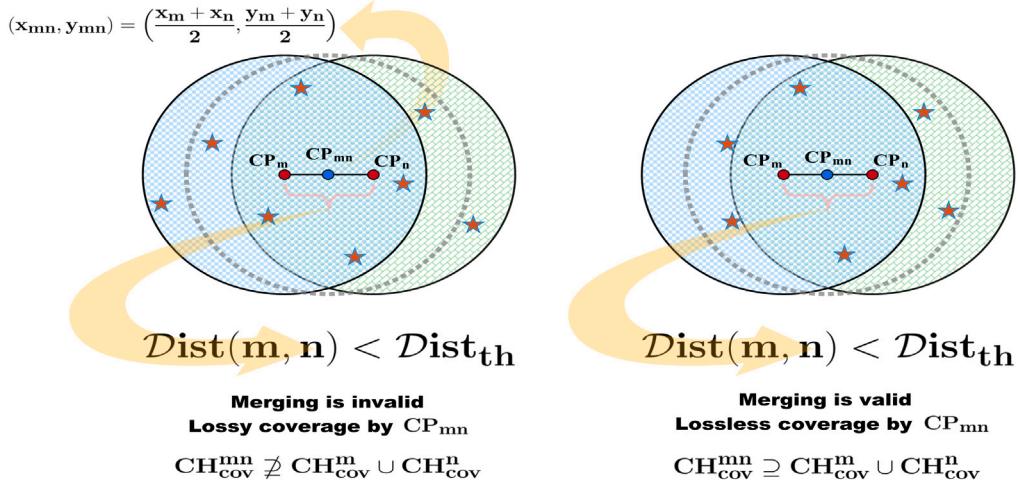


Fig. 4. Validity of merging.

coordinate values are either greater than  $(L-1)$  or less than 1. However, the proposed methodology requires the location coordinates to remain within the specified range. This issue can be resolved if we choose the updated positions as follows:

- If the updated  $x$  or  $y$  coordinate is less than 1 then replace the new coordinate with a randomly generated number in the range  $[1, 1 + \epsilon]$  where  $\epsilon$  is a small positive constant.
- If the updated  $x$  or  $y$  coordinate value is greater than  $(L-1)$  then replace the new coordinate with a randomly generated number in the range  $[(L-1 - \epsilon), (L-1)]$ .

#### 5.1.4. Dimensionality pruning

During the updation process, it has been observed that the positions (components) of every particle gradually approach to the global best

position. This eventually causes some of the particle components to increasingly get closer to each other. In other words, there can be several CPs in a small area that cover the redundant CH set. Moreover, the goal of the proposed protocol is to minimize the number of CPs on the MDC trajectory while satisfying the foregoing QoS requisites. Hence, it is required to prune the irrelevant components (dimension) of the higher dimensional particles in the swarm. This stimulates the proposed protocol to introduce the notion of **Dimensionality pruning** where the particles having greater dimensionality than  $g_{best}$  are considered. However, the shorter length particles in this context are not taken into consideration as they may explore a better solution in the multi-dimensional search space.

This strategy tells us that in higher dimensional particles, if the distance between any two CPs  $CP_m$  and  $CP_n$  is lesser than a predetermined threshold distance  $Dist_{th}$ , they might be considered for merging

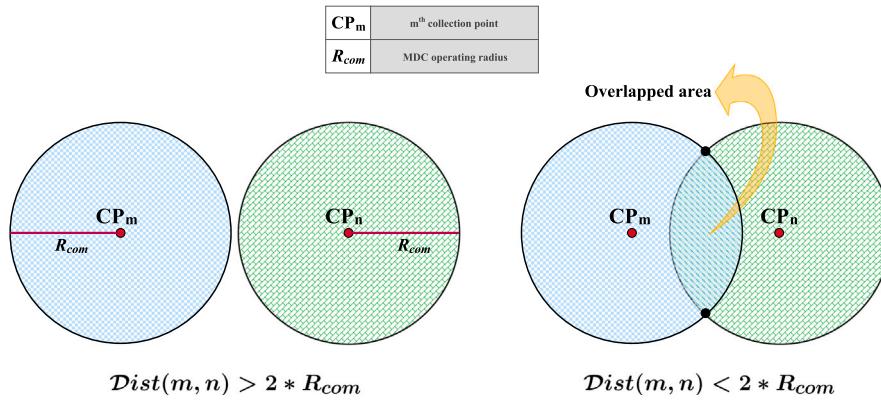
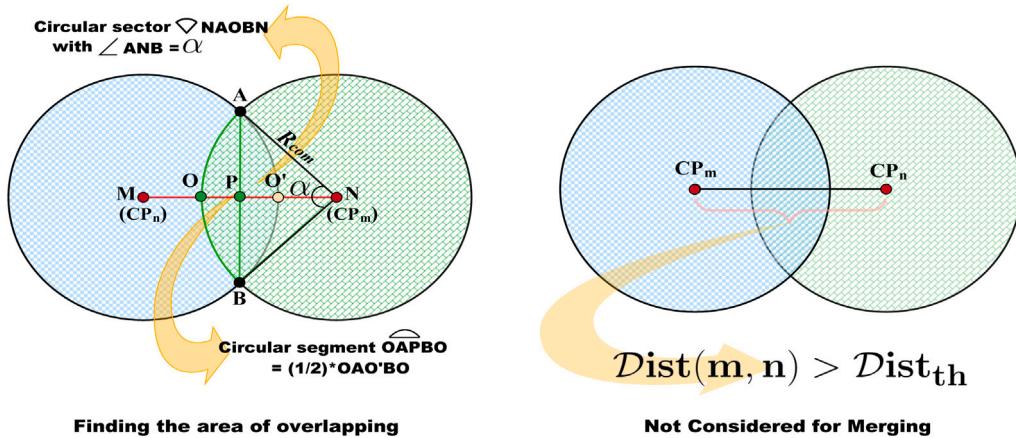


Fig. 5. MDC collection points at different distances.

Fig. 6. Estimation of overlapping area for  $Dist_{th}$ .

to produce a potential position  $CP_{mn}$ . The potential position  $CP_{mn}$  is formed by taking the midpoint of the line segment joining these proximate CPs. The position  $CP_{mn}$  is validated for merging  $CP_m$  and  $CP_n$  provided that a lossless coverage of the CHs is achieved (as shown in Fig. 4), i.e.,

$$\begin{cases} CH_{cov}^m \cup CH_{cov}^n \subseteq CH_{cov}^{mn} & \text{merging valid} \\ CH_{cov}^m \cup CH_{cov}^n \not\subseteq CH_{cov}^{mn} & \text{merging not valid} \end{cases}$$

where  $CH_{cov}^m$ ,  $CH_{cov}^n$ , and  $CH_{cov}^{mn}$  are the set of CHs covered by  $CP_m$ ,  $CP_n$ , and  $CP_{mn}$ , respectively. In other words, around  $CP_{mn}$ , if the MDC operating region (OR) with radius  $R_{com}$  can cover all the CHs that are originally covered by  $CP_m$  and  $CP_n$  then it leads to a valid merging, otherwise it is invalid.

The estimation of  $Dist_{th}$  depends on how much area is overlapped due to the intersection of ORs at  $CP_m$  and  $CP_n$ . It is obvious from Fig. 5 that two ORs overlap only when  $Dist(m, n) < 2 * R_{com}$  where  $Dist(m, n)$  denotes the distance between  $CP_m$  and  $CP_n$ . Herein,  $Dist_{th}$  value is considered as  $Dist(m, n)$  that produces a overlapping area of  $\left(\frac{1}{4}\right)^{th}$  size of the union of two ORs,  $OR_m$  and  $OR_n$ , i.e.,

$$Dist_{th} = Dist(m, n) \quad (16)$$

subject to  $\text{AREA}(OR_m \cap OR_n) = \frac{1}{4} \times \text{AREA}(OR_m \cup OR_n)$

Fig. 6 gives a better insight into overlapping area estimation of two proximate CPs where  $\text{AREA}(OR_m \cap OR_n) = \text{AREA}(OAO'BO) = 2 \times \text{AREA}(\text{Circular segment } OAPBO)$ . The following expressions are clearly visible from the mentioned figure:

$$\text{AREA}(OAPBO) = \text{AREA}(\text{Circular sector } \diamond NAOBN)$$

$$= \text{AREA}(\triangle ABN)$$

where the area of  $\diamond NAOBN$  and  $\triangle ABN$  can be calculated as:

$$\text{AREA}(\diamond NAOBC) = \pi R_{com}^2 \times \frac{\alpha^\circ}{360^\circ} \quad ; \quad \text{AREA}(\triangle ABC) = \frac{R_{com}^2 \times \sin \alpha}{2}$$

where  $\alpha$  is the central angle measured in degree and  $R_{com}$  is the MDC operating radius. Hence, area of the circular segment is derived as:

$$\text{AREA}(OAPBO) = \frac{R_{com}^2}{2} \times \left( \frac{\pi \alpha}{180} - \sin \alpha \right) \quad (17)$$

From Eqs. (16) and (17) the following equality can be derived as:

$$\begin{aligned} \text{AREA}(OAO'BO) &= \frac{1}{4} \times \text{AREA}(OR_m \cup OR_n) \\ \Rightarrow 2 \times \left[ \frac{R_{com}^2}{2} \times \left( \frac{\pi \alpha}{180} - \sin \alpha \right) \right] &= \frac{1}{4} \times (2\pi R_{com}^2 - \text{AREA}(OAO'BO)) \\ \Rightarrow R_{com}^2 \times \left( \frac{\pi \alpha}{180} - \sin \alpha \right) &= \frac{1}{4} \times \left( 2\pi R_{com}^2 - R_{com}^2 \times \left( \frac{\pi \alpha}{180} - \sin \alpha \right) \right) \\ \Rightarrow \frac{R_{com}^2 \pi \alpha}{180} - R_{com}^2 \sin \alpha &= \frac{2\pi R_{com}^2}{5} \\ \Rightarrow \sin \alpha &= \frac{\pi \alpha}{180} - \frac{2\pi}{5} \\ \Rightarrow 2 \times \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} &= \frac{\pi \alpha}{180} - \frac{2\pi}{5} \\ \Rightarrow 2 \times \sin \frac{\alpha}{2} \times \frac{NP}{R_{com}} &= \frac{\pi \alpha}{180} - \frac{2\pi}{5} \quad [\text{as shown in Fig. 6}] \\ \Rightarrow 2 \times NP &= \frac{8\pi R_{com}}{15\sqrt{3}} \quad [\text{central angle } \alpha \text{ is estimated as } 120^\circ] \\ \Rightarrow MN &\approx 0.97 R_{com} \end{aligned}$$

From the above calculation it has been derived that  $Dist_{th}$  is equals to  $0.97 * R_{com}$  (for the sake of simplicity, herein we consider the  $Dist_{th}$  value as  $R_{com}$ ).

In the next time step ( $t + 1$ ), again the fitness score of every particle will be evaluated as stated in Section 5.1.2 and based on that score each particle updates its  $pbest$  and the swarm updates the  $gbest$ . The updating process is defined below:

$$pbest_j^{t+1} = \begin{cases} X_j^{t+1} & \text{if } Fit(X_j^{t+1}) < Fit(pbest_j^t) \\ pbest_j^t & \text{otherwise} \end{cases} \quad (18)$$

$$gbest^{t+1} = \begin{cases} X_g^{t+1} & \text{if } Fit(X_g^{t+1}) < Fit(gbest^t) \\ gbest^t & \text{otherwise} \end{cases} \quad (19)$$

where  $X_g^{t+1}$  is the best position found in the whole swarm at ( $t + 1$ )th iteration. Then particle updation is accomplished using Eqs. (14), (15) which is followed by the dimensionality pruning (Section 5.1.4). This process continues until the iteration number reaches its predefined maximum value  $t_{max}$ . The final  $gbest$  vector is served as the optimal solution of CPs. Fig. 7 illustrates the working flow of the proposed VD-PSO algorithm in detail.

## 5.2. MDC trajectory construction

In order to gather cluster data, the second phase of the proposed protocol builds an optimal trajectory of the MDC that visits every CP exactly once and returns back to the BS by following a certain sequence. It is noteworthy that such a travelling sequence needs to generate the shortest length trajectory among all possible sequences to meet the delay-sensitive criteria of WSN. Now, for  $n_{cp}$  number of CPs, there will be  $\frac{(n_{cp}-1)!}{2}$  number of route possibilities as per the exhaustive search approach. However, this value will grow gigantically in case of a large-scale network; accordingly, the MDC trajectory construction experiences a non-polynomial time solution. Hence, it is highly desirable to utilize a construction based optimization algorithm like ACO to construct a near-optimal MDC trajectory in polynomial time. The essential parameters of ACO are given below:

- **Transition probability( $P_{ij}^k(t)$ ):** It refers to the probability of how ant  $k$  will choose CP  $j$  while sitting at CP  $i$  at time  $t$

$$P_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{p \in allowed_k} [\tau_{ip}(t)]^\alpha [\eta_{ip}]^\beta} & \text{if } j \in allowed_k \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

where  $allowed_k$  is the set of CP(s) which is (are) not yet visited by ant  $k$ ,  $\tau_{ij}(t)$  is the intensity of pheromone trail between the collection points  $i$  and  $j$  at time  $t$ ,  $\eta_{ij}$  is visibility of CP  $j$  from CP  $i$  i.e.,  $\eta_{ij} = \frac{i}{d_{ij}}$ , and,  $\alpha$  and  $\beta$  are parameters to control the relative influence of trail versus visibility.

- **Pheromone updating ( $\tau_{ij}(t + n_{cp})$ ):** Each ant completes a tour after  $n_{cp}$  iterations. Therefore, the pheromone trail  $\tau_{ij}(t + n_{cp})$  on edge  $e(i, j)$  at time  $t + n_{cp}$  can be updated as

$$\tau_{ij}(t + n_{cp}) = \rho \cdot \tau_{ij}(t) + \Delta \tau_{ij} \quad (21)$$

where  $\tau_{ij}(t)$  is pheromone trail on edge  $(i, j)$  at time  $t$ ,  $\rho$  is evaporation factor which regulates pheromone reduction, and,  $\Delta \tau_{ij}$  is total change in pheromone trail between time  $t$  and  $t + n_{cp}$  as shown below

$$\Delta \tau_{ij} = \sum_{k=1}^l \Delta \tau_{ij}^k \quad (22)$$

where  $l$  is the no of ants and  $\Delta \tau_{ij}^k$  is quantity per unit length of trail on edge  $e(i, j)$  by  $k$ th ant between time  $t$  and  $t + n_{cp}$

$$\Delta \tau_{ij}^k = \begin{cases} Q/L_k & \text{if ant } k \text{ travels on edge } e(i, j) \text{ between } t \text{ and } t + n_{cp} \\ 0 & \text{otherwise} \end{cases}$$

where  $Q$  is constant and  $L_k$  is tour length by ant  $k$ .

---

### Algorithm 1: MDC Trajectory Construction using ACO

---

**Data:** A fully connected weighted graph  $G = (\mathcal{V}, \mathcal{E})$   
**Result:** An optimal MDC tour sequence

```

1 begin
2    $t \leftarrow 0$  // time counter
3    $min\_list = []$ ;  $min\_len = []$  // store shortest tour and length
   in each iteration
4   for each edge  $e(i, j) \in E$  do
5      $\tau_{ij}(t) \leftarrow c$  where  $c > 0$ ;  $\Delta \tau_{ij} \leftarrow 0$ 
6   end
7   for  $NC = 1$  to  $NC_{max}$  do // cycle counter
8     Place  $l$  ants on  $n_{cp}$  number of CPs
9     for  $k = 1$  to  $l$  do
10       $List_k = []$  // for each ant  $k$  initialize  $List_k$ 
11       $List_k.append(start_k)$  // insert the starting CP of ant
          $k$  to the first index
12    end
13    for  $k = 1$  to  $l$  do
14      for  $cp = 1$  to  $(n_{cp} - 1)$  do // number of unvisited
         collection points
15        Move ant  $k$  to CP  $j$  based on  $P_{ij}^k(t)$  using Eq. (20)
16         $List_k.append(j)$  // insert the visited CP to the
         next index
17      end
18      Calculate  $L_k$ 
19    end
20    Find  $min_k L_k$ ,  $k = 1, 2, \dots, l$  and respective  $List_k$ 
21     $min\_len.append(L_k)$ ;  $min\_list.append(List_k)$ 
22    for each edge  $e(i, j) \in E$  do
23      Update  $\Delta \tau_{ij}$  using Eq. (22)
24      Compute  $\tau_{ij}(t + n_{sp})$  using Eq. (21)
25    end
26    if  $min\_len[NC] == min\_len[NC - 1]$  then
27      break
28    end
29  end
30  Print  $min\_list[NC]$ 
31 end

```

---

Algorithm 1 describes the working of ACO that produces a minimal length MDC trajectory from a fully connected weighted graph  $G = (\mathcal{V}, \mathcal{E})$ . Here,  $\mathcal{V}$  is the set of all CPs obtained by the proposed VD-PSO flowchart as shown in Fig. 7 and  $\mathcal{E}$  is the set of all possible edges among these CPs. The weight of an edge  $e(i, j) \in \mathcal{E}$  is measured by the euclidean distance between  $CP_i$  and  $CP_j$ . In Line 3 of Algorithm 1, two lists are defined that contain the shortest MDC trajectory and its length, respectively. Line 5 assigns a constant value  $c$  to the initial trail intensity  $\tau_{ij}(0)$  for each edge  $e(i, j) \in \mathcal{E}$ . Initially, a set of ants ( $l$ ) are randomly positioned over  $n_{cp}$  number of CPs. In Line 10, for each ant  $k$  a list  $List_k$  is defined to store the CPs already visited by ant  $k$ . Line 15 determines the movement of  $k$ th ant to the to CP  $j$  while sitting at CP  $i$  using the transition probability function defined in Eq. (20). For each ant  $k$  list  $List_k$  is updated by inserting the CP visited by that ant. Likewise, the set  $allowed_k$  can be updated as  $allowed_k = \{\mathcal{V} - List_k\}$ . Once, ant  $k$  completes a tour, the length of the MDC trajectory  $L_k$  is calculated. Among all  $l$  tours the minimal length tour and its length are stored in lists  $min\_list$  and  $min\_len$ , respectively. The process of updating

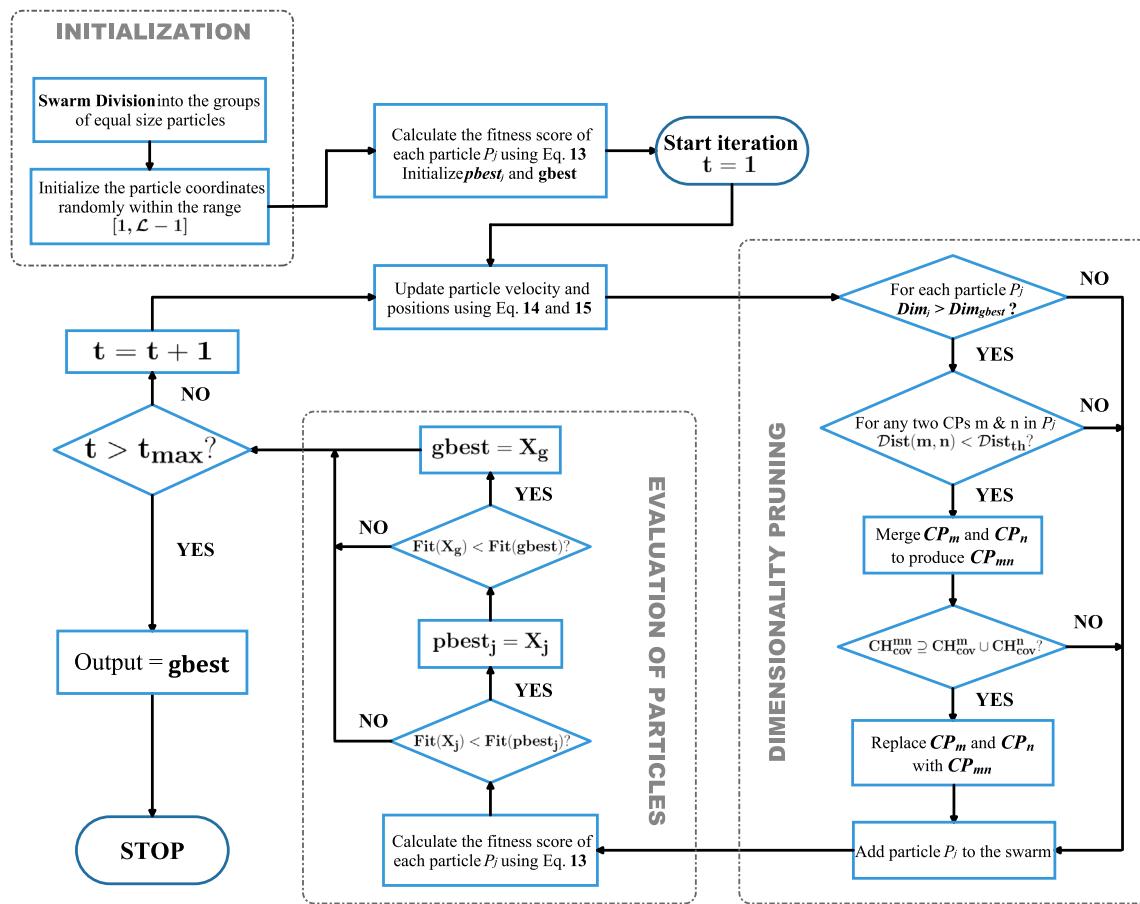


Fig. 7. Flowchart of the proposed VD-PSO algorithm.

the pheromone trail on edge is given by Lines 23 and 24. This iterative process continues till the maximum iteration value is reached or the solution converges (Line 26).

### 5.3. Steady phase

The data transmission process in WSN takes place in a bottom-up manner, i.e., it begins with the member nodes and ends with the CPs via intermediate CHs. For each cluster, the member nodes forward their sensory data to the respective CH following the time division multiple access schedule. Afterwards, data aggregation is carried out at each CH level to reduce data redundancy. In order to avoid the collision while multiple CHs simultaneously perform data forwarding towards the MDC, Code Division Multiple Access (CDMA) is employed by the WSN. The data gathering process is initiated by the MDC that sojourns at each CP by following the trajectory constructed in Algorithm 1. Upon arriving at each CP, MDC broadcasts a beacon message to wake up the neighbouring CHs within its operating radius  $R_{com}$ . To sojourn at each CP, a pause time is provided for the MDC that refers to the upper bound of the data gathering duration from the neighbouring CHs. Once the pause time expires, the MDC travels to the subsequent CPs before forwarding the accumulated data to the BS.

## 6. Simulation analysis

In this section, a comprehensive simulation analysis is conducted to manifest the superiority of the proposed protocol over the related existing protocols like PSO-ECSM [38], RBGM [31], and ACO-MSPD [40]. The whole set of simulations is executed in the development environment Spyder 3.1.2 using Python programming language. A list of

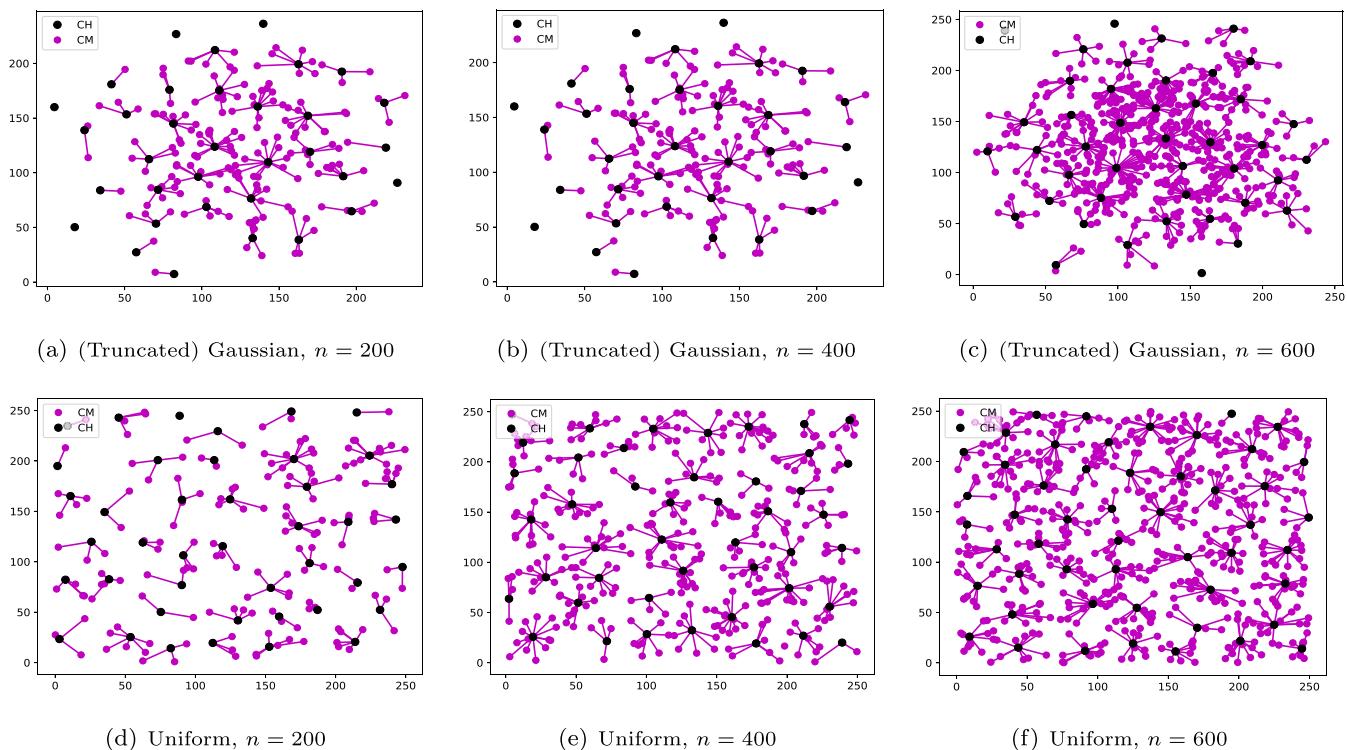
**Table 2**  
Simulation parameters.

Parameter	Value
Target region size ( $A$ )	250 × 250 m <sup>2</sup>
Number of nodes ( $n$ )	200–600
Initial energy of sensor nodes ( $E_{init}$ )	0.5 J
MDC operating radius ( $R_{com}$ )	56 m
Data packet size ( $k_d$ )	4000 bits
Control packet size ( $k_{cp}$ )	200 bits
Energy dissipation in the electronic circuit ( $\epsilon_{elec}$ )	50 nJ
Free space energy dissipation in the amplifier ( $\epsilon_{fs}$ )	10 pJ/bit/m <sup>2</sup>
Multi path energy dissipation in the amplifier ( $\epsilon_{mp}$ )	0.0013 pJ/bit/m <sup>4</sup>
Energy dissipation factor for data aggregation ( $\epsilon_{da}$ )	5 nJ/bit/signal

simulation parameters along with their values same as in [14] is shown in Table 2 whereas the parameters regarding the VD-PSO and ACO algorithms are presented in Tables 3 and 4 respectively. The proposed VD-PSO algorithm considers a swarm comprising 70 particles of variable lengths between 6–12 dimensions. The particle dimensionality is set as 15%–30% of the total cluster heads in the network. In the fitness function (Eq. (13)), the weight factor values are set as  $W_1 = 0.1$ ,  $W_2 = 0.3$ ,  $W_3 = 0.4$ ,  $W_4 = 0.2$ . Furthermore, a maximum number of 300 iterations ( $t_{max}$ ) is taken into account for the termination of the iterative process. Herein, we run each experimental setting 50 times in order to obtain a steady output and their average is shown as a final result.

### 6.1. Simulation environment

For simulation purpose, this article considers two distinct sensor deployments following uniform and truncated Gaussian distributions,



**Fig. 8.** Clustered networks with varying node densities for (truncated) Gaussian & uniform distributions.

**Table 3**  
VD-PSO parameters.

Parameter	Value
Swarm size ( $N_s$ )	70
Number of swarm divisions ( $N_g$ )	7
Max. number of iterations ( $t_{max}$ )	300
Particle lengths in dimension	6-12
Cognitive and social weights ( $c_1$ and $c_2$ )	1.495
Inertia weight ( $\omega$ )	0.5

**Table 4**  
ACO parameters.

Parameter	Value
Trail relative importance ( $\alpha$ )	1
Visibility relative importance ( $\beta$ )	1
Pheromone evaporation factor ( $\rho$ )	0.5
Constant related to the quantity of trail ( $Q$ )	100

respectively, to demonstrate the adaptability of the proposed protocol under various WSN applications. Uniformly distributed WSN, for instance, is suitable for continuous monitoring applications like battlefield surveillance that requires an immediate detection of the malicious intruders. Nevertheless, certain defence applications demand higher detection assurance at various locations in the deployment region. For instance, to defend a central station containing colonel and other high-rank officers from the enemies, it is necessary to deploy a greater number of sensor nodes around the station, while other non-critical items ought to be sparsely encircled by sensor nodes. This sort of deployment abides by the Gaussian distribution [47]. However, in traditional Gaussian distribution, some sensor nodes may be deployed outside the target region when the standard deviation exceeds some extent. In order to attain a bounded sensor deployment inside the target region, truncated Gaussian distribution [47] is considered herein. The mathematical notations of the respective probability density function (pdf)  $f(x, y)$  for the aforementioned distributions are presented below:

- A two-dimensional uniform distribution in the range  $[c, d]$  has a pdf as:

$$f(x, y) = \begin{cases} \frac{1}{d-c} & \text{for } c \leq x, y \leq d \\ 0 & \text{for } x, y < c \text{ or } x, y > d \end{cases} \quad (23)$$

where  $c = 0$  and  $d = 250$  are taken for the simulation purpose.

- A two-dimensional truncated Gaussian distribution with equal mean ( $\mu_x = \mu_y$ ) and equal standard deviation ( $\sigma_x = \sigma_y$ ) in both the dimensions (X and Y) of the target region has a pdf as:

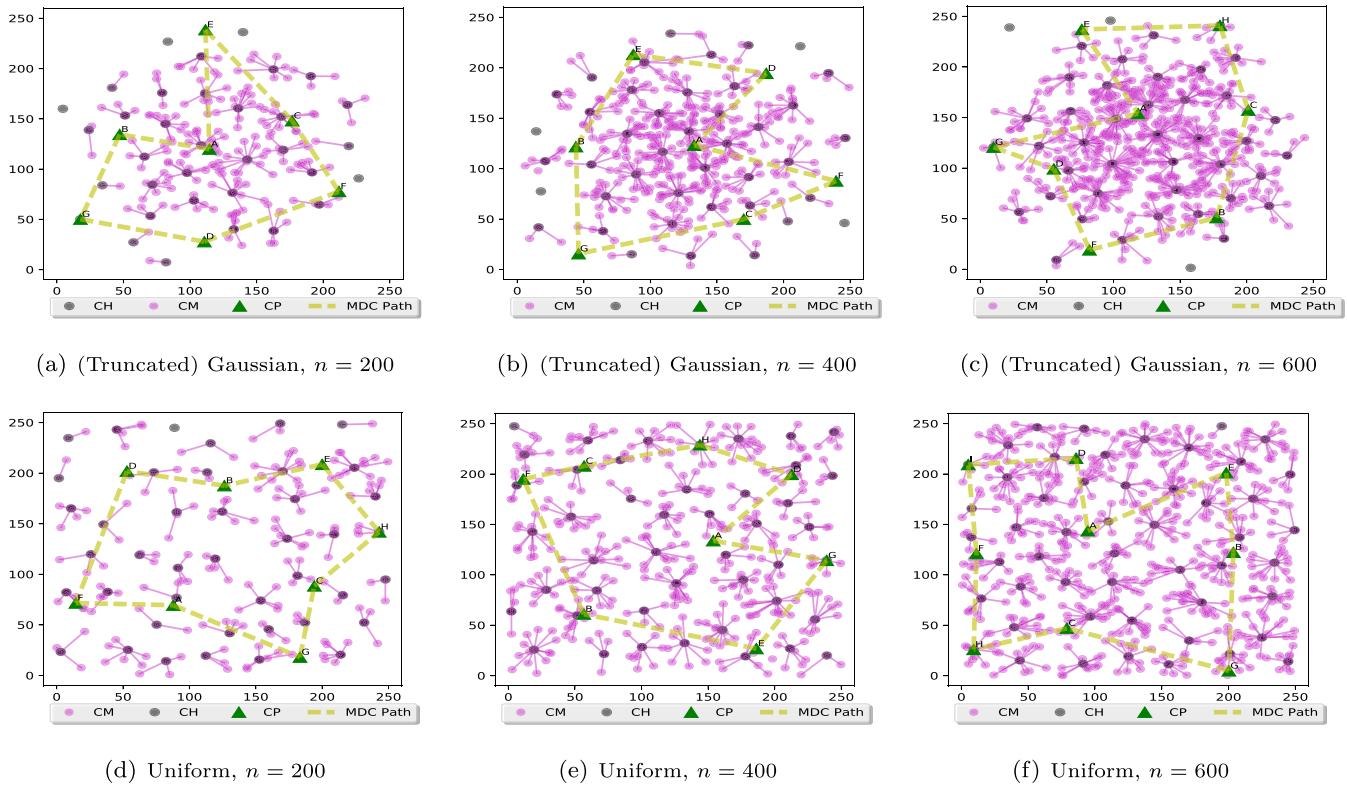
$$f^T(x, y, \sigma) = \frac{f(x, y, \sigma)}{\int_0^L \int_0^L f(x, y, \sigma) dy dx} \quad (24)$$

where  $L$  is the length of the target region side and. For simulation purpose, the parameter values for the two-dimensional truncated Gaussian distribution are taken as  $\mu_x = \mu_y = 125$ ,  $\sigma_x = \sigma_y = 50$ , and  $L = 250$ . In addition, is observed that varying the density of the sensor nodes could have a vital effect on the performance metrics of the network. Therefore, this article manifests the better scalability measure of the proposed protocol over the compared ones by conducting simulation analysis with the different number of nodes for both (truncated) Gaussian and uniform-distributed WSNs. Figs. 8(a)–8(c) exhibit clustered network scenarios under (truncated) Gaussian distributed WSN for 200, 400, and 600 number of nodes and 8(d)–8(f) show the same for uniformly distributed WSN. Here, the CH selection is obtained by following the clustering algorithm proposed in article [25]. Figs. 9(a)–9(f) exhibit the respective MDC trajectories for gathering cluster data from an optimal set of CPs.

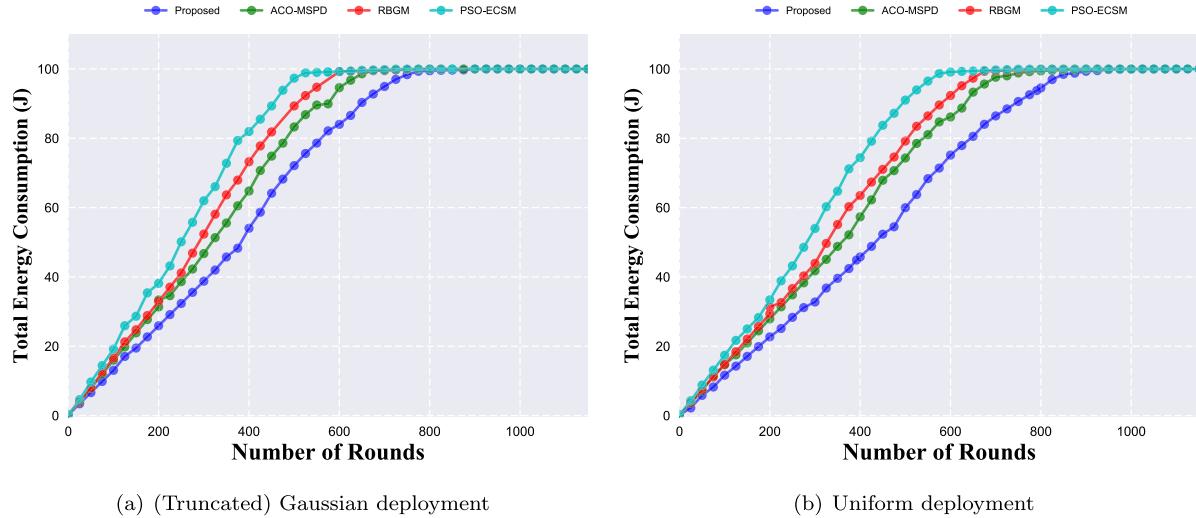
## 6.2. Benchmarking

For the benchmarking purpose, the following protocols are considered in this article:

- PSO-based Energy efficient Clustering and Sink Mobility (PSO-ECSM) algorithm [38]: This research aims to optimize the clustering as well as the sink mobility by exploiting the PSO algorithm.



**Fig. 9.** Optimal MDC trajectory for (truncated) Gaussian & uniform-distributed WSNs with varying densities.



**Fig. 10.** Total energy consumption where  $n = 200$ .

The fitness function for energy-efficient CH election integrates five parameters, namely node's residual energy, average energy, distance with the sink, number of neighbours, and energy consumption rate. On the other hand, sink mobility fitness function considers CH's residual energy, distance with the sink, and the cluster size.

- Routing Based on Grid structure and Mobile sink (RBGM) [31]: In this research, a virtual cellular infrastructure based routing protocol is designed that provides efficient data dissemination method between the sensor nodes and the mobile sink. Here, each cell member forwards its data to the respective cell header,

which in turn delivers them to the mobile sink. During the sink movement, the routing between cell headers and the sink is optimally updated with minimal delay and energy expenditure.

- ACO based Mobile Sink Path Determination (ACO-MSPD) algorithm [40]: The objective of this chain based algorithm is to evaluate an optimal set of Rendezvous Points (RPs) as well as sink trajectory under uneven data constraints. The nomination of optimal RPs begins with the formation of the directed spanning tree that counts the forwarding load for each sensor node. Afterwards, the ACO algorithm is leveraged to establish an optimal sink trajectory that passes through each RP.

The major motivation behind selecting these state-of-the-art protocols is that they cover a wide range of MDC based data gathering techniques such as cluster based and chain based. Considering this variety, we make an experimental comparison between the proposed protocol and them in terms of the following performance metrics.

### 6.3. Performance metrics

Based on the following performance metrics, the efficiency of the proposed protocol is evaluated.

#### 6.3.1. Energy efficiency

An energy-saving network is always desirable in energy-constrained applications of WSN because it ensures an extended network lifetime. Hence, MDC based WSN protocols should aim at minimizing the consumed energy in the network. Fig. 10 exhibits the competence of the proposed routing protocol in terms of total energy expenditure for 200 number of deployed sensors. It is evident that the proposed protocol surpasses the related existing protocols under both Gaussian and uniformly distributed scenarios. The better energy minimization is achieved in our model due to the construction of optimal MDC trajectory that leads to an energy-aware data gathering scheme. On the other hand, ACO-MSPD [40], and RBGM [31] despite attaining considerable energy minimization, encounter a multi-hop data gathering load. Moreover, RBGM [31] imposes an extra burden on the final cell headers since they need to deliver the whole network information to the MDC. This causes uneven energy consumption of the sensor nodes during data communication. In PSO-ECSM [38], an efficient sink mobility pattern is designed with the help of multi-objective particle swarm optimization (PSO). However, it does not guarantee the nomination of an optimal number of MDC sojourn locations which results in higher transmission energy dissipation. It is worth mentioning that the proposed protocol attains better energy efficiency under uniformly distributed scenarios than Gaussian distributed scenarios as Gaussian distribution encounters substantial variation in node density across the deployment region. This results in load unbalancing among the CHs, which leads to higher energy exhaustion of the network than the uniform ones. In addition, to evaluate the energy hole issue we present Table 5 that brings a statistical comparison among the residual energy population of the proposed and compared protocols for 100th, 200th, and 400th rounds. The column of “mean” shows the average residual energy of the sensor nodes measured in joule. Likewise, the column of “RSD” refers to the Relative Standard Deviation (RSD) of all nodes’ residual energy [25] and defined as:

$$RSD = \frac{\text{standard deviation of residual energy population } (\sigma)}{\text{mean of residual energy population } (\mu)} \times 100 \quad (25)$$

where lower value of RSD indicates lower deviation of the remaining energy population from the mean value ( $\mu$ ) i.e., better uniformness of the dissipated energy. It is clear from Table Table 5 that the proposed protocol is more energy balancing, and consequently guarantees better elimination of energy hole problem.

#### 6.3.2. Network lifetime

The utmost concern of any WSN application is to prolong the lifetime of the network. In this regard, Fig. 11 presents comparative line graphs for the studied protocols in terms of the number of alive nodes. The line graphs are plotted under Gaussian and uniformly distributed networks, respectively, with a 200 number of sensor nodes. It is obvious that the proposed protocol achieves major improvements over the compared protocols in extending the lifetime of the network. Enhancement of network lifetime is obtained due to the least energy consumption in each round caused by the free-space data transmission to the CPs. This article sets the terminology for network lifetime as First Node Die (FND) [25]. However, majority of the WSN protocols operate satisfactorily until half of the sensor nodes deplete their energy

(HND). Consequently, this paper measures the network lifetime of the studied protocols at two instances, i.e., FND and HND. Fig. 12 presents bar graphs that illustrate the dominance of the proposed method over the others in terms of both FND and HND for different node densities (200 – 600). From Fig. 12(a) it is observed that the proposed protocol improves the network lifetime with respect to FND by 20% – 30%  $\times$  than ACO-MSPD, 29% – 51%  $\times$  than RBGM, and 48% – 69%  $\times$  than PSO-ECSM protocol. Improvement of the network lifetime with respect to HND is achieved by 14% – 16%  $\times$  than ACO-MSPD, 20% – 25%  $\times$  than RBGM, and 24% – 29%  $\times$  than PSO-ECSM. From Fig. 12(b) we can notice that our protocol enhances the network lifetime with respect to FND by 23% – 39%  $\times$  than ACO-MSPD, 30% – 54%  $\times$  than RBGM, and 44% – 75%  $\times$  than PSO-ECSM protocol. Enhancement with respect to HND is obtained by 13% – 15%  $\times$  than ACO-MSPD, 18% – 20%  $\times$  than RBGM, and 27% – 31%  $\times$  than PSO-ECSM. The consistent improvements of the proposed protocol over the related protocols in all considered scenarios authenticates its better scalability. It should be noted that the proposed protocol gains higher network lifetime under uniformly distributed scenarios than the Gaussian ones as the former scenario leads to a better energy reduction of the network.

#### 6.3.3. Packet delivery ratio

In addition to the energy preservation of the network, QoS metrics are also essential for mission-critical, delay-sensitive, real-time multimedia, and many other WSN applications. Therefore, this study further investigates the performance of the studied protocol in terms of QoS routing metrics. As mentioned earlier, Packet Delivery Ratio (PDR) (Eq. (7)) is one of the key requisites of QoS routing that relies on the rate of congestion, type of the communication medium, type of the environment, and so on. The value of PDR defines the quality of link end-to-end reliability, where greater PDR value denotes better reliability. According to the random uniformed model [23], the probability of data packet loss dynamically increases with the increase in the transmission distance between source  $a$  and destination  $b$  and is defined by

$$P_{loss}(a, b) = \begin{cases} 0 & dist(a, b) \in [0, 50] \\ 0.01 * (dist(a, b) - 50) & dist(a, b) \in [50, 100] \\ 1 & dist(a, b) \in (100, \infty) \end{cases} \quad (26)$$

This indicates that a packet is assumed to be reached successfully when the link probability between the source and destination is greater than  $P_{loss}(a, b)$  otherwise, it is dropped. Figs. 13 and 14 provide insights into packet delivery ratio of all protocols by presenting box plots for various node densities (200 – 600) under both Gaussian and uniformly distributed deployments. Boxplot is a useful graph that describes the data distribution shape, its centrality, and variability with the help of three quartiles (Q1, Q2 or median, and Q3) and the range of the data. Fig. 13(a) manifests the superiority of the proposed protocol over the existing ones in terms of PDR till FND by 12% – 14%  $\times$  than ACO-MSPD [40], 14% – 16%  $\times$  than RBGM [31], and 22% – 24%  $\times$  than PSO-ECSM [38] while Fig. 14(a) shows the same till HND by 12% – 14%  $\times$  than ACO-MSPD, 17% – 19%  $\times$  than RBGM, and 24% – 29%  $\times$  than PSO-ECSM. Likewise, Fig. 13(b) shows that the proposed VD-PSO protocol increases the PDR till FND by 9% – 12%  $\times$  than ACO-MSPD, 14% – 17%  $\times$  than RBGM, and 21% – 26%  $\times$  than PSO-ECSM while Fig. 14(b) shows the same till HND by 11% – 16%  $\times$  than ACO-MSPD, 14% – 21%  $\times$  than RBGM, and 17% – 35%  $\times$  than PSO-ECSM. The superior performance is mainly contributed by minimizing the data transmission distance between CHs and CPs. On the contrary, ACO-MSPD and RBGM encounter multi-hop data transmission during data delivery that inflicts higher packet drop probability. Moreover, in RBGM, a severe burden of entire network data on the final cell header further enhances the packet drop. Likewise, PSO-ECSM does not provide one-hop data gathering strategy as it does not ensure an optimal number of CP generation due to the exploitation of traditional PSO. It should be mentioned

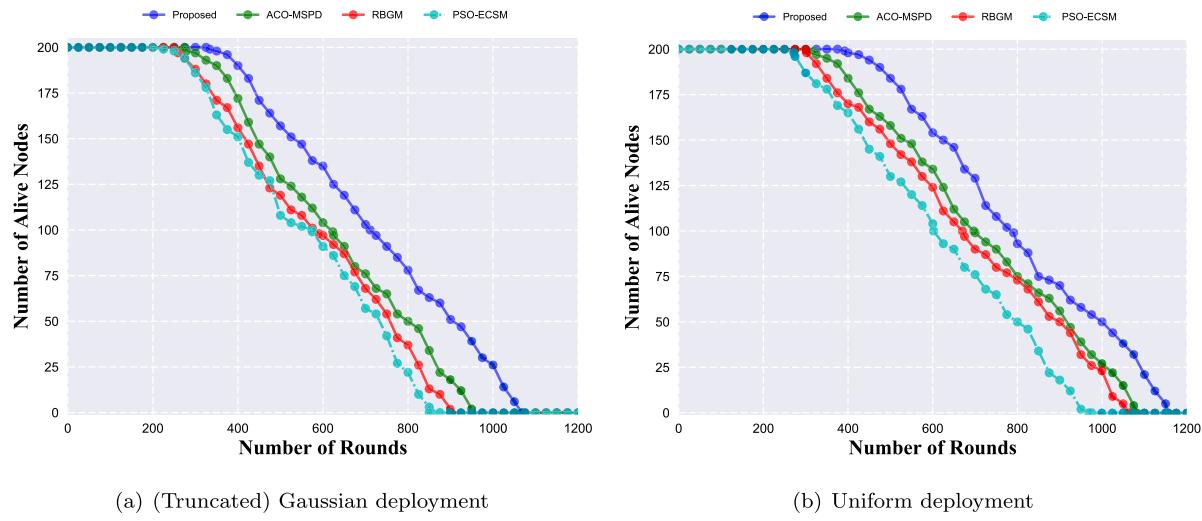
Fig. 11. Number of alive nodes where  $n = 200$ .

Table 5

Mean (joule) and relative standard deviation (%) of residual energy population for different rounds under various network scenarios.

Number of rounds	Protocol	Gaussian distributed WSN						Uniformly distributed WSN					
		200 nodes		400 nodes		600 nodes		200 nodes		400 nodes		600 nodes	
		mean	RSD (%)	mean	RSD (%)	mean	RSD (%)	mean	RSD (%)	mean	RSD (%)	mean	RSD (%)
100	PSO-ECSM	0.404	9.87	0.396	10.23	0.388	10.56	0.413	9.01	0.401	9.64	0.395	10.12
	RBGM	0.417	8.49	0.405	9.34	0.393	9.93	0.426	8.23	0.412	9.05	0.401	9.51
	ACO-MSPD	0.420	7.87	0.412	8.78	0.403	9.24	0.429	7.16	0.418	8.37	0.406	8.64
	Proposed	0.435	6.24	0.427	7.03	0.415	7.19	0.442	5.88	0.430	6.01	0.420	6.94
200	PSO-ECSM	0.309	18.70	0.298	19.64	0.290	20.37	0.333	17.54	0.324	19.26	0.315	19.67
	RBGM	0.335	17.20	0.326	18.64	0.317	18.97	0.353	16.66	0.344	17.86	0.333	18.13
	ACO-MSPD	0.343	14.56	0.332	16.23	0.320	16.85	0.360	13.81	0.351	14.84	0.342	15.60
	Proposed	0.370	11.53	0.358	12.68	0.345	13.65	0.386	11.47	0.374	12.31	0.366	12.78
400	PSO-ECSM	0.134	33.13	0.128	35.60	0.120	36.41	0.165	32.42	0.157	33.64	0.148	34.31
	RBGM	0.171	29.94	0.160	31.11	0.152	31.84	0.215	29.31	0.197	30.16	0.182	30.98
	ACO-MSPD	0.204	25.50	0.192	27.12	0.184	27.61	0.232	25.21	0.211	26.31	0.195	26.94
	Proposed	0.242	23.56	0.230	25.07	0.216	25.64	0.274	22.89	0.256	23.77	0.240	24.25

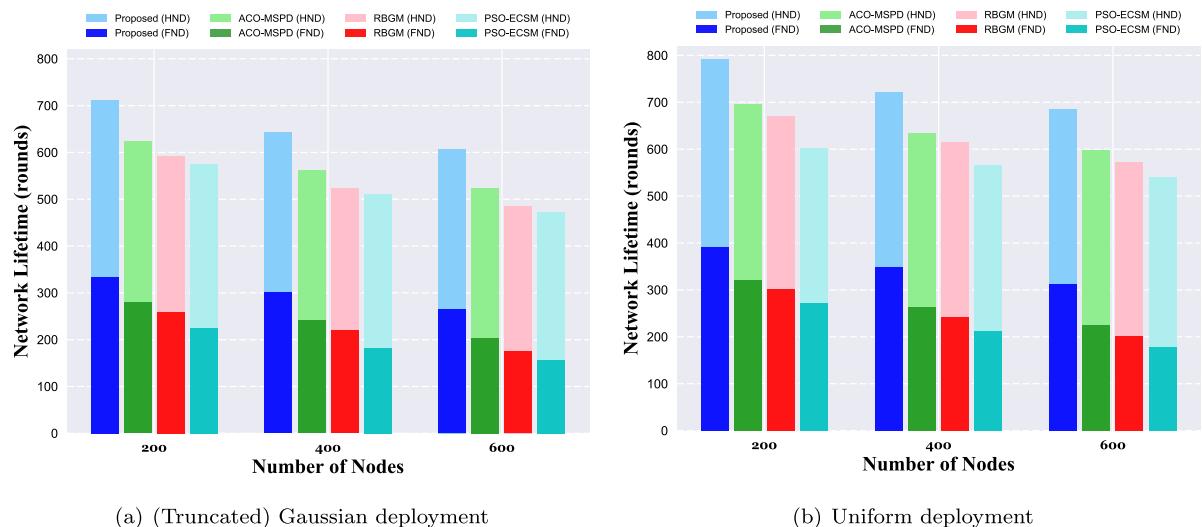


Fig. 12. Network lifetime.

that under uniformly distributed deployments, our protocol acquires higher PDR than the Gaussian distributed deployments. In Gaussian distributed networks, higher node density is observed around the centre

of the deployment region that leads to an uneven deployment of the sensor nodes. Moreover, during data delivery the probability of packet

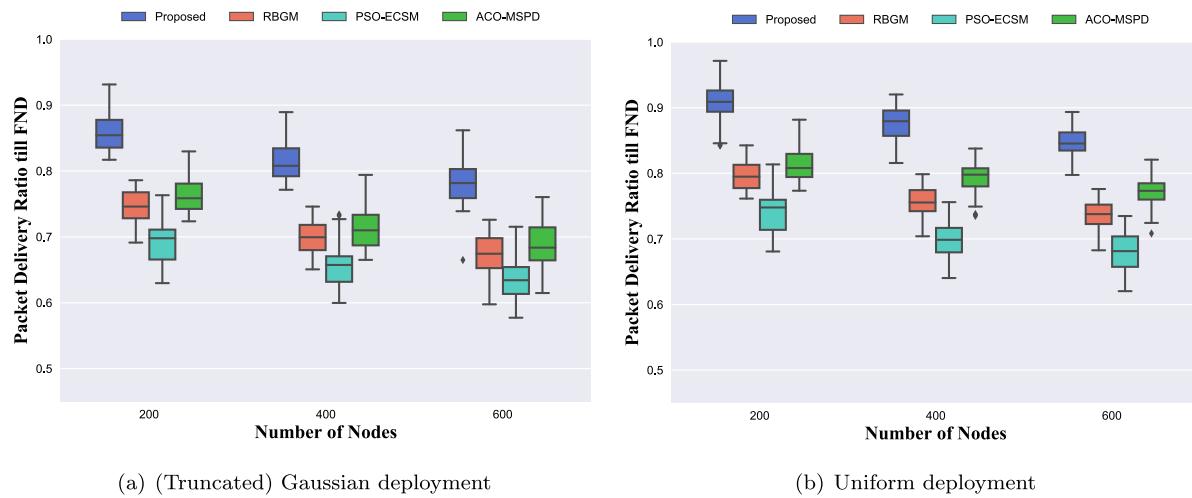


Fig. 13. Packet delivery ratio till first node die.

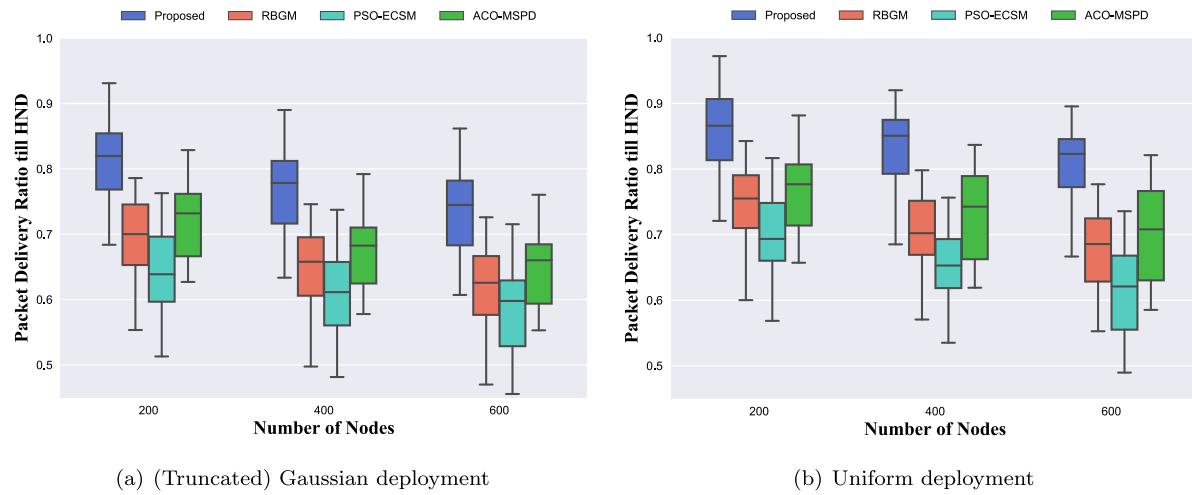


Fig. 14. Packet delivery ratio till half node die.

collision increases with the increase in node density; accordingly, it reduces the overall packet delivery of the network under Gaussian distribution.

#### 6.3.4. End-to-end delay

Another fundamental QoS routing metric is end-to-end delay ( $D_{e2e}$ ) of the transmitted data packets. As previously defined in Eq. (10) it is the time needed for the successful delivery of the generated data packets to the MDC. In delay-sensitive WSN applications, a lower value of  $D_{e2e}$  is always covetable since it implies the timeliness of the data delivery. Figs. 15 and 16 exhibit the superiority of the proposed VD-PSO protocol with respect to  $D_{e2e}$  till FND and HND, respectively by presenting box plots for various deployment scenarios. Fig. 15(a) exhibits that the proposed protocol reduces  $D_{e2e}$  value till FND by 11% – 13% × than ACO-MSPD [40], 16% – 19% × than RBGM [31], and 23% – 26% × than PSO-ECSM [38]. Likewise, Fig. 16(a) exhibits that the proposed one achieves lower  $D_{e2e}$  till HND by 10% – 12% × than ACO-MSPD, 17% – 24% × than RBGM, and 23% – 26% × than PSO-ECSM. Moreover, Fig. 15(b) shows the dominance of the proposed protocol in terms of  $D_{e2e}$  till FND by 12% – 14% × than ACO-MSPD, 20% – 22% × than RBGM, and 26% – 31% × than PSO-ECSM while Fig. 16(b) shows the same till HND by 13% – 15% × than ACO-MSPD, 20% – 22% × than RBGM, and 28% – 30% × than PSO-ECSM. The main reason for this is we propose an efficient data gathering strategy that reduces the data transmission distance between CHs and CPs. Moreover, reduction in MDC tour delay

is caused by building a shortest length MDC trajectory which minimizes the queuing delay. Conversely, ACO-MSPD and RBGM cause higher data delivery delay as they do not ensure a significant reduction in the data transmission distance. PSO-ECSM on the other side, due to the fixed-length particles, elongates the mobile sink traversal path that enhances the queuing delay. It is noteworthy that the end-to-end delay tends to increase with the increase in node density of the network. This happens because a higher number of nodes increases network traffic, which significantly degrades the data delivery timeliness.

#### 6.4. Analysis of VD-PSO algorithm

Herein, we present the adaptive nature of the proposed VD-PSO algorithm for various deployment scenarios. Fig. 17(a) exhibits the optimal number of CPs generated by the proposed algorithm for different node densities (200 – 600) under both Gaussian and uniformly distributed WSN. It should be noted that the number of elected CPs for the same population of nodes is different for different node deployments. This highlights the importance of adopting a flexible optimization algorithm like VD-PSO. Moreover, the quantity of CPs increases monotonically with the increase in node densities to balance the data gathering load among the CHs. An increasing number of CPs eventually elongates the MDC trajectory length (in meter) as shown in Fig. 17(b) for all considered scenarios.

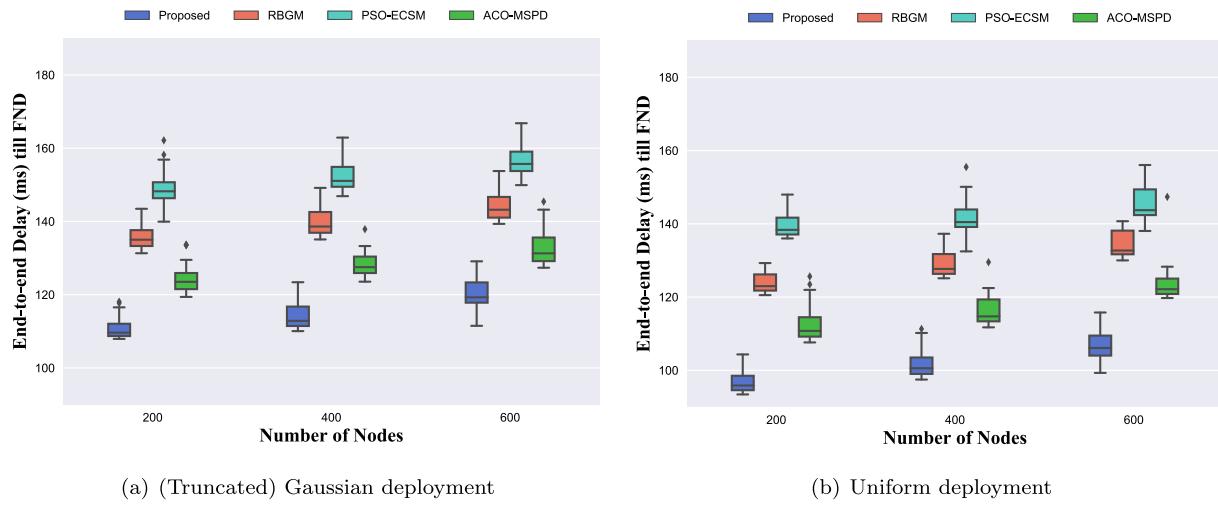


Fig. 15. End-to-end delay till first node die.

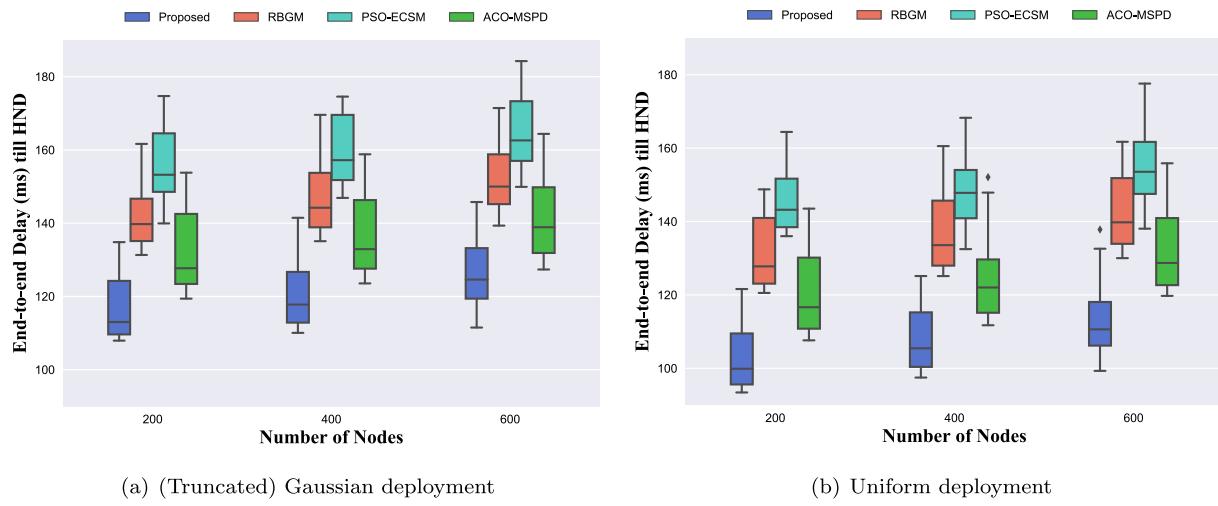


Fig. 16. End-to-end delay till half node die.

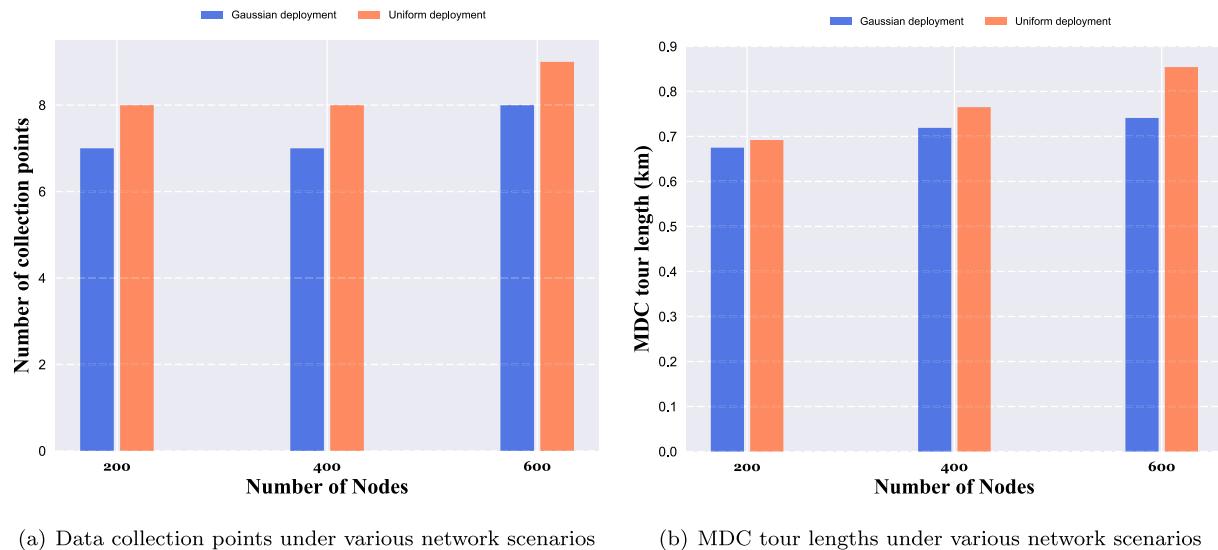
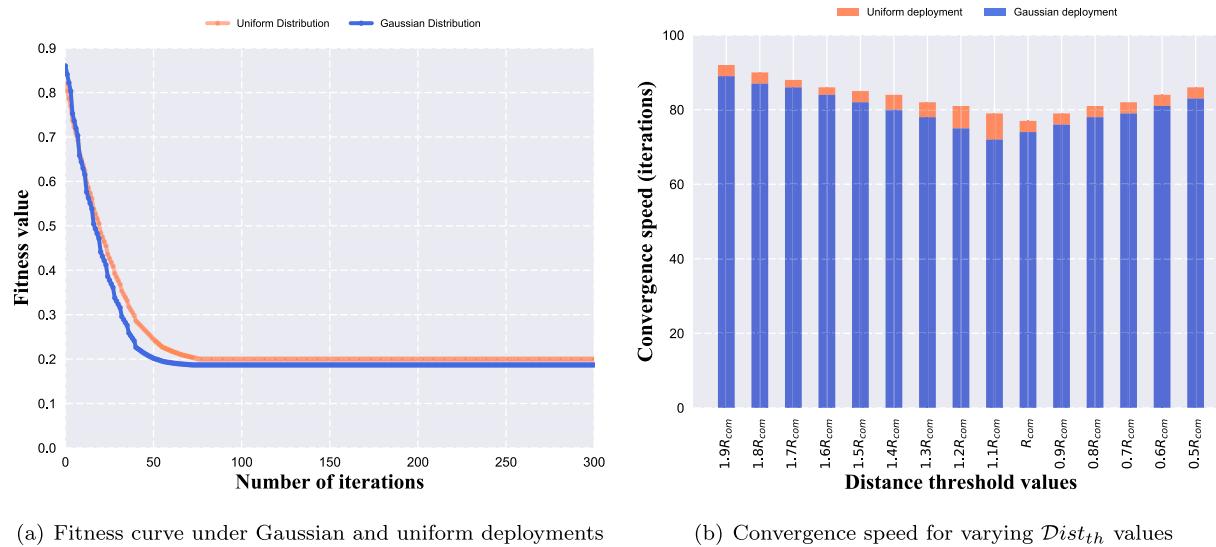


Fig. 17. Adaptability of the proposed VD-PSO algorithm.



**Fig. 18.** Fitness function curve and the rate of convergence for 200 number of nodes.

As stated earlier, each feasible MDC trajectory in the proposed VD-PSO algorithm maps to a particle of CPs whose fitness score is estimated with the help of a multi-objective fitness function. Fig. 18(a) depicts the fitness curve of the proposed VD-PSO algorithm for 200 number of sensor nodes that refers to the swarm's global best in each iteration. It should be noted that in this graph, we consider an average of the gbest values produced over 50 runs. Furthermore, it is observed that the proposed algorithm attains a global optimal solution for both Gaussian and uniform deployments. In Fig. 18(b), we analyse different values of  $Dist_{th}$  to verify the optimal convergence rate of the proposed optimization algorithm. It can be observed that the best convergence rate is achieved at  $Dist_{th} = R_{com}$  which validates the estimated  $Dist_{th}$  value in sub Section 5.1.4.

## 7. Conclusion

In this article, we have investigated the problem of discovering an optimal set of CPs such that each CP is surrounded by the maximum possible number of CHs in one-hop proximity. In order to nominate such an optimal set of CPs, we have proposed an MDC based data gathering protocol that provides an effective balance between energy consumption and data gathering delay in the network. Our protocol leverages the VD-PSO algorithm to optimize the location and number of the SPs. VD-PSO improves the SP nomination by commencing its iterative search process with different dimensional particles. The quality of each particle is assessed by applying a QoS parameter based fitness function. Such multi-objective fitness function design confirms that the achieved solution form a fine tuning among the multiple conflicting parameters. Moreover, improved particle updation and dimensionality pruning methods are also introduced to speed up the convergence of the VD-PSO algorithm. The optimal CP nomination not only improves the QoS parameters of the WSN but also limits the data transmission energy depletion. In order to form an optimal trajectory of the MDC for visiting the CPs, the Ant Colony Optimization algorithm is incorporated. The overall performance of the proposed algorithm is compared with some related existing schemes named PSO-ECSM, RBGM, and ACO-MSPD. The simulated results demonstrate its efficiency for several QoS parameters like energy efficiency, network lifetime, end-to-end delay, packet delivery ratio, etc.

Nevertheless, this research considers the terrestrial motion of MDC, which can be extended to aerial motion in the light of UAV-based IoT applications. In addition, UAVs typically have resource constraints,

such as buffer and power. Therefore, this work can be expanded to include applications in multi-constraint UAV based data collection.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] Ruth Ande, Bamidele Adebisi, Mohammad Hammoudeh, Jibran Saleem, Internet of Things: Evolution and technologies from a security perspective, Sustainable Cities Soc. 54 (2020) 101728.
- [2] Pradeep K Khatua, Vigna K Ramachandaramurthy, Padmanathan Kasinathan, Jia Ying Yong, Jagadeesh Pasupuleti, Arul Rajagopalan, Application and assessment of internet of things toward the sustainability of energy systems: Challenges and issues, Sustainable Cities Soc. 53 (2020) 101957.
- [3] Anagha Rajput, Vinod Babu Kumaravelu, Scalable and sustainable wireless sensor networks for agricultural application of Internet of things using fuzzy c-means algorithm, Sustain. Comput.: Inform. Syst. 22 (2019) 62–74.
- [4] R.P. Meenaakshi Sundhari, K. Jaikumar, IoT assisted hierarchical computation Strategic Making (HCSM) and Dynamic Stochastic Optimization Technique (DSOT) for energy optimization in wireless sensor networks for smart city monitoring, Comput. Commun. 150 (2020) 226–234.
- [5] Jaspreet Singh, Ranjit Kaur, Damanpreet Singh, A survey and taxonomy on energy management schemes in wireless sensor networks, J. Syst. Archit. (2020) 101782.
- [6] Bushra Rashid, Mubashir Husain Rehmani, Applications of wireless sensor networks for urban areas: A survey, J. Netw. Comput. Appl. 60 (2016) 192–219.
- [7] R. Annie Uthra, S.V. Kasmir Raja, Qos routing in wireless sensor networks—A survey, ACM Comput. Surv. 45 (1) (2012) 1–12.
- [8] Babar Nazir, Halabi Hasbullah, Energy efficient and QoS aware routing protocol for clustered wireless sensor network, Comput. Electr. Eng. 39 (8) (2013) 2425–2441.
- [9] Habib Mostafaei, Energy-efficient algorithm for reliable routing of wireless sensor networks, IEEE Trans. Ind. Electron. 66 (7) (2018) 5567–5575.
- [10] Tarunpreet Kaur, Dilip Kumar, A survey on QoS mechanisms in WSN for computational intelligence based routing protocols, Wirel. Netw. 26 (4) (2020) 2465–2486.
- [11] Nabajyoti Mazumdar, Amitava Nag, Sukumar Nandi, HDDS: Hierarchical Data Dissemination Strategy for energy optimization in dynamic wireless sensor network under harsh environments, Ad Hoc Netw. 111, 102348.
- [12] Amin Shahrazi, Amir Taherkordi, Øystein Haugen, Frank Eliassen, Clustering objectives in wireless sensor networks: A survey and research direction analysis, Comput. Netw. (2020) 107376.
- [13] Muhammad Aslam, Ehsan Ullah Munir, M Mustafa Rafique, Xiaopeng Hu, Adaptive energy-efficient clustering path planning routing protocols for heterogeneous wireless sensor networks, Sustain. Comput.: Inform. Syst. 12 (2016) 57–71.

- [14] Wendi Rabiner Heinzelman, Anantha Chandrakasan, Hari Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, in: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, IEEE, 2000, pp. 10–pp.
- [15] Sara Al-Sodairi, Ridha Ouni, Reliable and energy-efficient multi-hop LEACH-based clustering protocol for wireless sensor networks, *Sustain. Comput.: Inform. Syst.* 20 (2018) 1–13.
- [16] Deepak Mehta, Sharad Saxena, MCH-EOR: Multi-objective cluster head based energy-aware optimized routing algorithm in wireless sensor networks, *Sustain. Comput.: Inform. Syst.* (2020) 100406.
- [17] Ju Ren, Yaxue Zhang, Kuan Zhang, Anfeng Liu, Jianer Chen, Xuemin Sherman Shen, Lifetime and energy hole evolution analysis in data-gathering wireless sensor networks, *IEEE Trans. Ind. Inf.* 12 (2) (2015) 788–800.
- [18] Reem E Mohamed, Ahmed I Saleh, Maher Abdelrazzak, Ahmed S Samra, Energy-efficient routing protocols for solving energy hole problem in wireless sensor networks, *Comput. Netw.* 114 (2017) 51–66.
- [19] Ying-Gao Yue, Ping He, A comprehensive survey on the reliability of mobile wireless sensor networks: Taxonomy, challenges, and future directions, *Inf. Fusion* 44 (2018) 188–204.
- [20] Jun Tao, Liang He, Yanyan Zhuang, Jianping Pan, Maryam Ahmadi, Sweeping and active skipping in wireless sensor networks with mobile elements, in: 2012 IEEE Global Communications Conference (GLOBECOM), IEEE, 2012, pp. 106–111.
- [21] Jiqiang Tang, Hongyu Huang, Songtao Guo, Yuanyuan Yang, Dellat: Delivery latency minimization in wireless sensor networks with mobile sink, *J. Parallel Distrib. Comput.* 83 (2015) 133–142.
- [22] Tzung-Shi Chen, Wei-Qing Du, Jen-Jee Chen, Energy-efficient data collection by mobile sink in wireless sensor networks, in: 2019 IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2019, pp. 1–6.
- [23] Mohammed Abo-Zahhad, Sabah M Ahmed, Nabil Sabor, Shigenobu Sasaki, Mobile sink-based adaptive immune energy-efficient clustering protocol for improving the lifetime and stability period of wireless sensor networks, *IEEE Sens. J.* 15 (8) (2015) 4576–4586.
- [24] Anjula Mehto, Shashikala Tapaswi, K.K. Pattanaik, Virtual grid-based rendezvous point and sojourn location selection for energy and delay efficient data acquisition in wireless sensor networks with mobile sink, *Wirel. Netw.* (2020) 1–17.
- [25] Saugata Roy, Nabajyoti Mazumdar, Rajendra Pamula, An energy and coverage sensitive approach to hierarchical data collection for mobile sink based wireless sensor networks, *J. Ambient. Intell. Humaniz. Comput.* (2020) 1–25.
- [26] Maryam Naghibi, Hamid Barati, EGRPM: Energy efficient geographic routing protocol based on mobile sink in wireless sensor networks, *Sustain. Comput.: Inform. Syst.* 25 (2020) 100377.
- [27] Ramin Yarinezhad, Reducing delay and prolonging the lifetime of wireless sensor network using efficient routing protocol based on mobile sink and virtual infrastructure, *Ad Hoc Netw.* 84 (2019) 42–55.
- [28] Weimin Wen, Shenghui Zhao, Cuijuan Shang, Chih-Yung Chang, EAPC: Energy-aware path construction for data collection using mobile sink in wireless sensor networks, *IEEE Sens. J.* 18 (2) (2017) 890–901.
- [29] Areej Alsafin, Ahmed M. Khedr, Zaher Al Aghbari, Distributed trajectory design for data gathering using mobile sink in wireless sensor networks, *AEU-Int. J. Electron. Commun.* 96 (2018) 1–12.
- [30] Ramin Yarinezhad, Seyyed Naser Hashemi, A cellular data dissemination model for wireless sensor networks, *Pervasive Mob. Comput.* 48 (2018) 118–136.
- [31] Ramin Yarinezhad, Seyyed Naser Hashemi, An efficient data dissemination model for wireless sensor networks, *Wirel. Netw.* 25 (6) (2019) 3419–3439.
- [32] Gao Ren, Juebo Wu, Frederik Verdonen, Bee-based reliable data collection for mobile wireless sensor network, *Cluster Comput.* 22 (4) (2019) 9251–9260.
- [33] Hui Wang, Kangshun Li, Witold Pedrycz, An elite hybrid metaheuristic optimization algorithm for maximizing wireless sensor networks lifetime with a sink node, *IEEE Sens. J.* 20 (10) (2020) 5634–5649.
- [34] Saunhita Sapre, S. Mini, A differential moth flame optimization algorithm for mobile sink trajectory, *Peer-to-Peer Netw. Appl.* (2020) 1–14.
- [35] Anjula Mehto, Shashikala Tapaswi, K.K. Pattanaik, Pso-based rendezvous point selection for delay efficient trajectory formation for mobile sink in wireless sensor networks, in: 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS), IEEE, 2020, pp. 252–258.
- [36] Amar Kaswan, Vishakha Singh, Prasanta K. Jana, A multi-objective and PSO based energy efficient path design for mobile sink in wireless sensor networks, *Pervasive Mob. Comput.* 46 (2018) 122–136.
- [37] Xiaolin He, Xiuwen Fu, Yongsheng Yang, Energy-efficient trajectory planning algorithm based on multi-objective PSO for the mobile sink in wireless sensor networks, *IEEE Access* 7 (2019) 176204–176217.
- [38] Biswa Mohan Sahoo, Tarachand Amgoth, Hari Mohan Pandey, Particle swarm optimization based energy efficient clustering and sink mobility in heterogeneous wireless sensor network, *Ad Hoc Netw.* (2020) 102237.
- [39] Muralitharan Krishnan, Sangwoon Yun, Yoon Mo Jung, Dynamic clustering approach with ACO-based mobile sink for data collection in WSNs, *Wirel. Netw.* 25 (8) (2019) 4859–4871.
- [40] Praveen Kumar, Tarachand Amgoth, Chandra Sekhara Rao Annavarapu, ACO-based mobile sink path determination for wireless sensor networks under non-uniform data constraints, *Appl. Soft Comput.* 69 (2018) 528–540.
- [41] Jin Wang, Jiayi Cao, R. Simon Sherratt, Jong Hyuk Park, An improved ant colony optimization-based approach with mobile sink for wireless sensor networks, *J. Supercomput.* 74 (12) (2018) 6633–6645.
- [42] Bassam Faiz Gumaida, Juan Luo, Novel localization algorithm for wireless sensor network based on intelligent water drops, *Wirel. Netw.* 25 (2) (2019) 597–609.
- [43] Zhaoyang Wang, Baihai Zhang, Xiaoyi Wang, Xuebo Jin, Yuting Bai, Improvements of multihop localization algorithm for wireless sensor networks, *IEEE Syst. J.* 13 (1) (2018) 365–376.
- [44] Wei Kuang Lai, Chung Shuo Fan, Lin Yan Lin, Arranging cluster sizes and transmission ranges for wireless sensor networks, *Inform. Sci.* 183 (1) (2012) 117–131.
- [45] Muralitharan Krishnan, Sangwoon Yun, Yoon Mo Jung, Improved clustering with firefly-optimization-based mobile data collector for wireless sensor networks, *AEU-Int. J. Electron. Commun.* 97 (2018) 242–251.
- [46] Chafika Benzaïd, Miloud Bagaa, Mohamed Younis, Efficient clock synchronization for clustered wireless sensor networks, *Ad Hoc Netw.* 56 (2017) 13–27.
- [47] Yun Wang, Weihuang Fu, Dharma P. Agrawal, Gaussian versus uniform distribution for intrusion detection in wireless sensor networks, *IEEE Trans. Parallel Distrib. Syst.* 24 (2) (2012) 342–355.



**Saugata Roy** received B.Tech. degree in Information Technology from Maulana Abul Kalam Azad University of Technology, West Bengal and M.Tech. degree in Computer Science & Engineering from Indian Institute of Technology (ISM), Dhanbad, India. Currently, he is pursuing his Ph.D. degree in the Department of Computer Science and Engineering at Indian Institute of Technology (ISM) Dhanbad, India. His research interest includes wireless sensor networks, internet of things, and soft computing.



**Nabajyoti Mazumdar** received B.E. degree in Information Technology from The University of Burdwan, India. He did his M.Tech. and Ph.D. degree in Computer Science and Engineering from Indian Institute of Technology (ISM), Dhanbad, India, in 2014 and 2018 respectively. Currently, he is an Assistant Professor in the Department of Information Technology at Indian Institute of Information Technology Allahabad, India. He has acted as referees in many reputed journals including IEEE Internet of Things, IEEE transactions on industrial informatics, IEEE System Journal, Simulation Modelling Practice and Theory, International Journal of Communication Systems, etc. His current research interest includes wireless sensor network, internet of things, and soft computing.



**Rajendra Pamula** received B.Tech. (CSE) degree from Rajiv Gandhi College of Engineering, Nandyal affiliated to Jawaharlal Nehru Technological University Hyderabad, India, in the year 2001 and M.Tech. (CSE) degree from Gandhi Institute of Technology and Management, Visakhapatnam affiliated to Andhra University, Visakhapatnam, India in the year 2004. He received his Ph.D. degree from Indian Institute of Technology, Guwahati, Assam, India, in the year 2015 under the supervision of Prof. Sukumar Nandi. At present he is working as an Assistant Professor in the Department of Computer Science and Engineering, Indian Institute of Technology (ISM) Dhanbad, India. He has more than 60 research publications in various national and international journals and conference proceedings. His current research interest includes data mining and computer networks.