

# Single Mobile Sink Based Energy Efficiency and Fast Data Gathering Protocol for Wireless Sensor Networks

Shivkumar S. Jawaligi<sup>1</sup>, G. S. Biradar<sup>2</sup>

<sup>1</sup>Electronics & Communication Engineering, APPA Institute of Engineering & Technology, Kalaburagi, India

<sup>2</sup>Electronics & Communication Engineering, PDA College of Engineering, Kalaburagi, India

Email: shiv.jawaligi@gmail.com, gsbiradar@gmail.com

**How to cite this paper:** Jawaligi, S.S. and Biradar, G.S. (2017) Single Mobile Sink Based Energy Efficiency and Fast Data Gathering Protocol for Wireless Sensor Networks. *Wireless Sensor Network*, 9, 117-144. <https://doi.org/10.4236/wsn.2017.94007>

**Received:** March 8, 2017

**Accepted:** April 25, 2017

**Published:** April 28, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Recently, the exponential rise in communication system demands has motivated global academia-industry to develop efficient communication technologies to fulfill energy efficiency and Quality of Service (QoS) demands. Wireless Sensor Network (WSN) being one of the most efficient technologies possesses immense potential to serve major communication purposes including civil, defense and industrial purposes etc. The inclusion of sensor-mobility with WSN has broadened application horizon. The effectiveness of WSNs can be characterized by its ability to perform efficient data gathering and transmission to the base station for decision process. Clustering based routing scheme has been one of the dominating techniques for WSN systems; however key issues like, cluster formation, selection of the number of clusters and cluster heads, and data transmission decision from sensors to the mobile sink have always been an open research area. In this paper, a robust and energy efficient single mobile sink based WSN data gathering protocol is proposed. Unlike existing approaches, an enhanced centralized clustering model is developed on the basis of expectation-maximization (EEM) concept. Further, it is strengthened by using an optimal cluster count estimation technique that ensures that the number of clusters in the network region doesn't introduce unwanted energy exhaustion. Meanwhile, the relative distance between sensor node and cluster head as well as mobile sink is used to make transmission (path) decision. Results exhibit that the proposed EEM based clustering with optimal cluster selection and optimal dynamic transmission decision enables higher throughput, fast data gathering, minima delay and energy consumption, and higher efficiency.

## Keywords

Wireless Sensor Network, Data Gathering, Single Mobile Sink Node, Centra-

## 1. Introduction

The high pace rise in the wireless communication system and associated communication applications have given rise to a new arena where across academia-industries the emphasis has been made to develop certain efficient and robust communication systems. In present day scenario, communication systems have been playing vital role across human presence by serving major purposes including civil utilities and surveillance applications, security systems, data communication, defence systems, business communications, industrial communication, and many more. In recent years, “Internet of Things” commonly known as IoT has gained recognition across academia-industries because of its immense potential to serve major communication purposes. However, exploiting this technology sternly demands high Quality of Service (QoS) provision and network efficiency with minimal cost, end-to-end delay, minimal computational overheads and energy consumption. To meet these demands enriching Wireless Sensor Networks (WSNs) has always been first choice. On the other hand, the decentralized and Ad-hoc nature of WSN make it frontrunner in industry for easy deployment and low cost infrastructure. In fact WSN is a communication network comprising multiple nodes operating simultaneously within certain expected region of interest, where it performs event-data collection in a well defined collaborative manner. The comprising nodes collect the data and forward it to the nearest neighbor node towards destination or sink where the collected data are used to make decision. The robustness of WSN makes it's a suitable alternative for low cost communication solution for major advanced communication systems, including IoT ecosystems, underwater sensor networks, civil/defense/industrial applications etc. Considering functional paradigm, WSN is a battery powered network where each node is equipped with a battery source. However, being functional in exhaustive network it undergoes continuous energy consumption that eventually reduces the network (node) lifetime. On the other hand maintaining optimal QoS provision is must for WSN based advanced communication networks. As a cumulative optimization measure reducing energy exhaustion and end-to-end delay, and maintain higher throughput and network efficiency is must for WSN. These factors motivate us to develop a more robust and efficient routing system for WSNs.

Exploring in depth of the various literatures, it can be easily found that data gathering process is the foundation of the WSN technologies. In general WSN system, data gathering at the sink node is done using generic static node and multi-hop communication paradigm. However, repeated data transmission and multi-hop traversal make the traditional approach inefficient, particularly energy exhaustive and delayed, which somewhere confines the employability for time critical communication applications. On the other hand, considering load sensi-

tive communication the multi-hop communication with static nodes is not suggested universally due to non-deniable data loss probability, latency and re-transmission caused energy exhaustion. Furthermore, in WSN the sink nodes collect the transmitted data from different sensor nodes (including source node as well as intermediate nodes) to make certain decision. This overall process exhibits significantly higher computations. On the contrary major delivering timely data with QoS assured communication is inevitable for present day communication need [1] [2].

Considering aforementioned demands, in last few years efforts have been made to introduce mobility feature in WSN so as to reduce multi-hop transmission and associated data drop probability. Introducing mobility with the WSN can greatly enhance the overall network performance by reducing unwanted multi-hop data transmission and traversal period thereby delivering timely data gathering without significant packet drop. However, introducing mobility in WSN makes system to undergo contention and hence data drop. A number of efforts are being made to enhance routing protocol to enable mobility based WSN for efficient data gathering [1]-[6]. However, there are a number of issues in mobile sink based data gathering in WSN. Some of the key issues in mobile sink based data gathering are presented as follows:

### Key Issues of Data Gathering in Mobile Sink Based WSN

- Being mobile the collecting nodes undergoes dynamic state conditions and it functions as an intermediate node so as to deliver data from the sensor node to the base station (BS). In this process, the dual traversal (*i.e.*, from sensor node to the mobile sink and from mobile sink to the BS) introduces delay that could confine the network efficiency in data gathering.
- In WSN, energy efficiency is the dominating issue. In case of more nodes in the network the energy exhaustion becomes more and in addition causes higher computational complexities (*i.e.*, node table management, signaling overheads, forwarding route selection etc).
- The simultaneous gathering of the large amount of data from multiple sensors can significantly minimize the energy exhaustion in WSN, which can be efficiently done using single mobile sink node. However, managing contention and collision in the network is tedious task due to greedy nature of the WSN nodes.
- There are many real time application scenarios, such as industrial application, vehicular communication, battlefield utilities etc, where timely data delivery is must. However, since the BS is far away from the mobile collector, it might violate delay constraints thus resulting into hazardous consequences.

In addition, there are numerous challenges in the traditional WSN data gathering schemes. Some of them are given as follows:

- In WSN, the data from the sensor node is routed to the BS on the basis of the network topology. On the contrary, due to mobility based network
- During data gathering the same path can be used by multiple sensors so as to

transfer their data to the BS that as a result can cause contention. However, the efficient routing model to deal with such contention can significantly reduce energy exhaustion and can avoid overheads due to alternate forwarding path selection.

- In practice multiple sensor nodes can transmit its data to the BS concurrently that as a result increase the probability of packet drop and retransmission causing energy exhaustion and delay.

No doubt, numerous efforts have been made to reduce the existing limitations of the data gathering protocols where mobile sink has been suggested to be used for data gathering, however exploring real time scenarios, it can be found that the predominant limitations of the mobile sink based data gathering is that the collection time to the sink can be higher than stationary sink based approach. This can be due to higher traveling time or traversal conditions across the network. During this traversal period the possibility of packet loss can't be ignored, particularly in that condition in which the mobile sink could not reach sender at the appropriate time. In present day scenario majority of applications, including major IoT applications are time-sensitive and hence require fast data gathering ability so as to make earlier decision. In addition, reducing the number of clusters is also must as eventually it too causes energy exhaustion. No doubt, increasing the transmission range (say communication radius) may lead reduction in the forwarding hop, however may trigger energy consumption by nodes. It signifies a trade-off between traversal period and associated network lifetime.

This research paper intends to exploit the significances and robustness of the clustering approach which performs grouping of the nodes on the basis of the geographical locations and forms certain clusters. This as a result avoids the need of forwarding nodes and hence reduces the channel load, contention and computational complexities. No doubt, clustering technique has always been a robust approach in which nodes exhibit Cluster Head (CH) selection so as to perform data transmission. In this approach, the nodes perform data transmission to CH, where the collected data is further transmitted to the BS. In this approach, only a few nodes transmit data to the BS located far away and hence it avoids energy exhaustion. In addition, it eliminates the issue of network traffic and hence contention probability. Unlike multiple mobile node based data gathering, this research paper intends to develop a robust single mobile sink based data gathering in WSN that could not only alleviate the contention issues but could make the overall communication efficiency optimal. Single mobile sink based WSN data gathering can significantly reduce the signaling overheads that as a result can make network computational efficient. In addition, it can avoid major issues such as getting trapped in any security related adversaries [1-2]. With these motivations, in this paper we intend to develop a robust single mobile sink node based data gathering protocol for WSN networks. Considering the key aspect of clustering based single mobile sink based data gathering, in this paper an enhanced data transmission model has been developed that at first employs robust heuristic approach to decide best node mobility path so as to enable

swift reach ability of the sink node to collect data from the CH. In generic approaches CH collects data from each sensor nodes and then forwards it to the sink (even mobile sink). However, the fact that a sensor node can directly transmit the data to the mobile sink (without transmitting through CH) when comes in proximity or the transmission range has not been addressed in any research or literature. In our work, we intend to explore this feasibility so as to enable fast data gathering to assist mission critical and QoS transmission. The performance of the proposed routing approach has been examined in terms of energy consumption, delay, network efficiency and throughput, where it has been found better than existing approaches.

The remaining section of this paper is divided as: Section II discusses the related works for mobile sink based data collection. The contribution of the presented study is given in section III, which is followed by the discussion of the experimental setup and results obtained in Section IV. Research conclusion and future scope is discussed in section V. References used are presented at the last of manuscript.

## 2. Related Work

To enable an efficient single mobile sink based data gathering in last few years a number of efforts have been made. In this section, some of the key literatures discussing single mobile sink based data gathering schemes are discussed.

In [7] a load-sensitive clustering approach was developed where authors incorporated cost metric and metric-mapping schemes to derive an energy efficient WSN data gathering protocol. To perform data gathering in WSN, authors [7] [8] developed mobility based routing protocol where they emphasized on enhancing the delay sensitive and energy efficient routing. Realizing the need of mission-critical data gathering in WSN authors [4] proposed a Query-Based Data Collection Scheme (QBDCS). Their approach incorporated each-cycle data gathering model, where they observed it efficient in assuring minimal end-to-end delay and energy consumption. However, authors could not assess the impact of iterative transmission and gathering complexity. As enhancement authors [9] derived a dual-phase data gathering model, in which first they employed network information gathering protocol (NIGP) to achieve network state conditions and node information and based on available information scheduled data forwarding to the sink node. Unfortunately, it couldn't deal with the issues of the computational overheads and end-to-end delay incurred because of multi-step procedure. In addition, they could not address the topological variation and its impact on node management to perform data gathering. In [10] distributed energy efficient adaptive clustering protocol (DEACP) was developed to perform WSN data gathering. Authors emphasized on developing an energy efficient and scalable network design. Considering mobile sink based data gathering in WSN, they applied self-controlled mobile sink, which exhibited data gathering in three phases; mobile sink movement, data gathering and node localization, and mobile sink localization. The application of self-controlled mobile

sink exploited node localization and data source identification to perform sink trajectory decision for data collection. In [11] a weight based realistic clustering algorithm (WRCA) was developed in which authors applied multiple mobile nodes to perform data gathering over Ad-hoc networks. Considering network security during data gathering, authors [3] developed detection avoidance model named transverse forward through (TFT). In [12] a traversal length-constrained mobile sink based data gathering approach was developed where they applied clustering approach to derive the best trajectory projections for data gathering [12]. A single mobile sink based data gathering model was developed in [13] in which at first the complete network was divided into multiple regions of equal size, which was then followed by clustering in each region based on the degree of freedom. Authors applied K-means clustering approach to perform cluster head (CH) selection which was further applied to perform data gathering from the connected cluster nodes (CN).

Some other works [14]-[19] too concluded that mobile sink based data gathering can be efficient to enable not only energy efficiency but also computational efficiency. Interestingly, author in [20] questioned over the justification provided [14]-[19]. Author [20] advocated that multiple sink based data gathering is more efficient than single mobile sink, particularly for delay sensitive communication. However, they [20] could not address the issue of energy exhaustion and computational complexity due to multiple mobile sink nodes. The issue of best forwarding node selection using integer linear programming for data gathering over WSN was discussed in [21]. At first they emphasized over reducing path planning cost followed by energy exhaustion minimization. In [22], authors compared the efficiency of the single mobile sink and multiple mobile sink based data gathering over WSN. In fact, with a distributed network condition the movement pattern of the mobile sink node has vital impact on network performance. Considering this fact authors in [22] incorporated random walk mobility model for which the next probable sink location was decided based on best forwarding node selection.

Considering robustness of the clustering approach authors [23] applied single mobile sink to derive a tree-cluster-based data-gathering over WSN. Authors introduced a weight-based tree-formation approach to perform data gathering. Considering data gathering of the large size authors [15] applied mobile sink node. However, they stated that with mobile sink based data gathering estimating optimal mobile sink trajectory and cluster formation are the tedious task to achieve energy efficient communication. In [24] a clustering based WSN data gathering protocol was developed where authors applied mobile sink to collect data from static sensor nodes. Authors [25] [26] suggested that mobility based data gathering can be of paramount significance, especially for large size sensor networks. In [27] developed an energy efficient data transmission model where they exploited the efficiency of multiple inputs and multiple output (MIMO) and Multi-hop transmission model along with clustering. In [28], moving vehicle based data gathering protocol was developed in which sensor nodes at first lo-

cate the mobile sink (vehicle) and forwards its data towards BS. They emphasized on higher throughput while maintaining minimum hop-counts. In comparison to the uncontrolled mobility sink based data gathering their proposed model exhibited better in terms of throughput and energy efficiency. Considering reliability during mobile sink node movement, authors developed a high-reliability data gathering protocol for WSN network [29]. They applied random mobile sink node movement where sensors transmit their data periodically. Their model exploited node information to and beacon message information to decide mobile sink path. To perform load balancing and next hop identification, authors focused on reducing the packet loss ratio caused due to node mobility. In addition, they applied adaptive beacon strategy and no-route buffer mechanism to perform data gathering. A simple mobile sink based data gathering model was developed in [30] in which mobile sink splits sensor nodes into grids irrespective of the location of the sensor nodes and employs random walk mobility pattern to perform patrolling. However, such type of clustering which is not based on the node position may cause data drop significantly and thus can reduce overall performance. In addition, it can't be stated to be time efficient due to random movement across network.

No doubt, Low-Energy Adaptive Clustering Hierarchy (LEACH) [31] has always been one of the efficient and dominating clustering algorithms for communication over WSNs using the static sink node. In this approach, clustering is performed by the individual sensor node and the sensor nodes perform information exchange such as residual energies and thus based on the node with higher residual energy is selected to be the CH. However, LEACH possesses numerous limitations. For illustration, because LEACH functionally assumes that the individual sensor node can communicate with other nodes in the network, the WSNs deployed in wide areas often fail to employ this approach. Majority of the distributed approaches such as LEACH generally employ the limitation of the communication range of a node in the network. Some of the key distributed clustering approaches developed so far are K-hop Overlapping Clustering Algorithm (KOCA) [32] and k-hop connectivity ID (k-CONID) [33]. The first approach (KOCA) is emphasized on multiple overlapping clusters and to perform CH selection and node location estimation applies probabilistic approach. On the contrary, in k- CONID routing scheme which is a probabilistic scheme the comprising nodes perform respective random IDs exchange with each other and the node possessing the minimum ID within k-hop is selected as a CH.

Considering load sensitive data gathering application over WSNs, reducing data transmission is highly intricate task, especially for a distributed clustering approach. In case a sensor network is split physically into small network subsets, it becomes difficult for a node to have entire network information. With intent to reduce energy consumption; in this paper centralized clustering paradigm (CCP) has been taken into consideration. Furthermore, the CCP, which is performed by a super node, is optimal for mobile sink based data gathering. Some of the well known centralized clustering algorithms are Power-efficient data ga-

thering in Sensor Information Systems (PEGASIS) [34] and K-means based on travelling salesman problem (TSP) for mobility based data gathering. Authors [34] formed node chain based clustering where they applied node location information to perform clustering. It considers the drawbacks of the radio range of the sensor nodes, and thereby ensures uniform energy consumption across network. However, it [34] does not exhibit even satisfactory (minimal) energy consumption due to greedy nature of the algorithm. On the contrary, authors in [35] employed mobility model in which it splits the sensor nodes into multiple clusters using k-means clustering approach. K-means clustering being a CCP model functional on the basis of the location of the nodes, gives clustering outcome very close to the optimization. This is the matter of fact that the resulting optimal clustering can reduce energy consumption, however its [35] mobility approach was designed without applying the communication range condition that as a result can force the mobile sink to drop data during collection from different nodes.

In general, most of the available approaches for the sensor node based clustering are categorized into three predominant types; centralized algorithms which is a distributed algorithm that does not consider any significant node information such as node location and its communication range, distributed algorithms that avoids any need of nodes' information, and the third approach is the distributed algorithms that applies nodes' location and communication range to perform network clustering. Interestingly, in order to accomplish a novel solution while maintaining minimal transmission and data gathering from all the nodes distributed across the WSN, there is the need of a novel centralized paradigm, employing key node characteristics such as its location and communication range. With such motivations, unlike major existing approaches in this paper the emphasis has been made on developing a clustering model with minimal transmission overheads to perform data gathering. In addition, here to enable fast data gathering where the sensor node can transmit its data to the CH as well as directly to the mobile sink on the basis of the radio range criteria.

### 3. Proposed Model

In this paper, a multi-objective target model has been derived that intends to exploit various efficient (feasible) approaches to perform data gathering in densely distributed WSNs. This research emphasizes on the following:

- Exploring the effectiveness of the mobile sink for delay sensitive, energy efficient data gathering,
- Examining the impact of data request (control messages) on the energy exhaustion,
- Estimating the optimal number of clusters to perform efficient data transmission,
- Introducing a responsiveness model to estimate best transmission decision among CH, mobile sink and the sensor nodes so as to enable fast (delay sensitive) data gathering,



- Assessing the performance effectiveness of the proposed routing scheme by comparing with other well known CCP based routing schemes.

In addition, in this research we intend to examine whether the increase in cluster counts increases data gathering efficiency by reducing data requests.

The discussion of the overall proposed research model is presented as follows:

### **Centralized Clustering Paradigm (CCP) Based Data Gathering in WSNs**

This section briefs about clustering problem in sensor network, data gathering using mobile sink and associated measures to achieve optimal performance. In addition, the network model being considered and the expectation maximization model and their relevancy significance towards efficient data gathering are also discussed in this section.

#### **Clustering**

This is the matter of fact that to develop a robust data gathering scheme for WSN using mobile sink, the predominant issue is to reduce energy consumption by estimating the optimal position where data gathering has to be performed.

In a simple manner, the overall problems can be stated as the answers for the following two questions:

- What should be the optimal approach to split nodes to form clusters?
- What should be the optimal number of clusters in the network so as to reduce energy consumption and end-to-end delay in data gathering?
- Can introducing additional scheduling constraint named “distance sensitive data transmission and forward Scheduling” between sensor nodes and the cluster heads or mobile sink node can enhance network performance in terms of energy consumption, delay, efficiency, throughput etc.?

Considering a generic concept that the needed energy for data transmission of node is often in relation (*i.e.*, proportional) to the square of the transmission distance, the optimal approach to reduce energy consumption (for data gathering) can be to reduce the sum of square (SoS) of transmission distance in WSN. To achieve these goals, recently an approach called Expectation Maximization (EM) was proposed which has exhibited optimal, particularly for solving the clustering problem by means of estimating mathematical function (say formulae) iteratively. Considering the fact that expectation maximization model can significantly reduce the SoS of the inter-node distance and the cluster head (CH), in our proposed model EM approach has been taken into consideration over the 2D Gaussian Mixture Model (GMM) distribution. On the contrary, it is also undeniable that in practical environment there exist certain limitations, particularly for the maximum communication range and not every sensor nodes in the network can communicate to each other, and even with the cluster centroid (here onwards we state it as cluster head or CH). In this case, those sensor nodes which are unable to communication with CH require communicating in multi-hop fashion. In case of multi-hop communication paradigm, the communica-

tion distance signifies the SoS in between the nodes coming across in the path (multi-hop distance to the BS) to the BS. It affirms the fact that the communication distance is often distinct as compared to the direct distance between a sensor node and BS. However, the expectation maximization approach intends to reduce the SoS of the direct distance, rather than reducing the communication distance. Therefore, it becomes inevitable to employ expectation maximization model, especially in that WSN environment where the maximum communication range is limited. EM incorporates enhancement in such manner that it reduces the SoS of the communication distance.

#### 4. WSN Network Model

To assess the effectiveness of the proposed routing model for data gathering, in this research work a WSN network comprising a mobile sink and multiple sensor nodes placed across the network within a limited field. In the proposed network model each node is aware of its location by means of localization model. In addition, the mobile sink node too knows the information (*i.e.*, location information) about all nodes distributed across the network. This is the fact that irrespective of being a sensor or sink, each node possesses fixed communication range  $R$  and hence the communication can be successful only within the communication range  $R$ . In the proposed network model, the mobile sink node performs patrols the CH estimated so as to reduce energy exhaustion introduced because of the data transmission, and thus performs data gathering from the sensor nodes. Here, each sensor node is equipped with a defined resource availability or buffer-size where it stores retrieved (*i.e.*, sensed) network information or the data information till mobile sink approaches the connected CH. Thus, the information collected at the individual node is transferred to the mobile sink in multi-hop approach. In addition, based on a new factor called responsiveness the sensor node decides whether it should transmit data through CH or directly to the mobile sink which is nearer than the CH in that network region. It makes data transmission decision more effective and time efficient.

Considering a practical communication scenario, in this paper WSN has been assumed as a densely distributed network that reflects urban communication scenarios, schools or any regionally spread-out network, boarder regions and mountains etc. Under such assumptions, the overall network is split into sub-network regions. **Figure 1** illustrates the considered network model. As stated in **Figure 2** circles signify the  $N$  sensor nodes distributed across the network region  $L \times L$ , where  $L$  presents geographical parameter. Here, the filled circle  $C$  signifies the cluster head which is supposed to be visited by the mobile sink node for data gathering. Here, the solid-fill region presented the group of the nodes, while dotted circle signifies the cluster. Noticeably, in our model the term “group” refers the set of nodes capable of communicating with each other. On the contrary, due to inter-node distance reason, the nodes belonging to different groups are unable to communicate with each other. Here, the variable  $G$  presents the number of groups in the network region, and the other variables

**Algorithm 1** Enhanced EM (EEM) clustering algorithm for WSN data gathering

**Input:** CHs ( $\sigma$ )

**Output:** CH ( $\sigma$ ), covariance matrix ( $\Sigma$ ), and the nodes belonging to the individual cluster

Initialize the input parameters: CHs, ( $\sigma$ ) and place at the arbitrary positions across the network.

Estimate clustering parameters ( $\sigma$ ) and  $\Sigma$ .

Calculate the parameters,  $D_{nc}$  and  $E$ .

**while**  $|E - E^{new}| < \gamma$  **do**

With the highest  $S_g$  value, perform "Group ( $g$ ) selection".

**for**  $c \in C_g$  **do**

**for**  $n \in N_g$  **do**

Estimate RF value ( $\varphi_{nc}$ ) for the  $n$ th node

**end for**

Estimate the total number of nodes  $N_c$  belonging to the cluster,

Update the parameters  $\sigma$ , and  $\Sigma$ , using  $N_c$ .

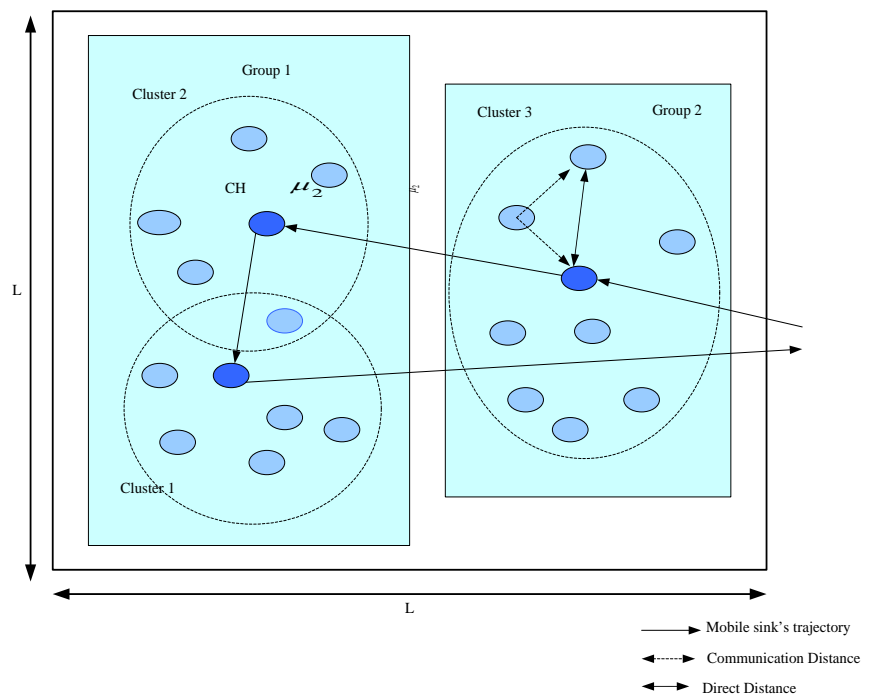
**end for**

Estimate log likelihood  $E^{new}$ .

**end while**

Return CH ( $\sigma$ ), covariance matrix ( $\Sigma$ ), and the nodes belonging to the individual cluster.

**Figure 1.** Enhanced EM (EEM) clustering algorithm for WSN data gathering.



**Figure 2.** An illustration of the considered WSN network.

$N_g$  and  $C_g$  signifies the total number of sensor nodes and the total number of clusters in the  $g$ th group, correspondingly. In our model, we have estimated the total number of groups by exploiting key information such as the location of the nodes and their respective communication range  $R$ . In our proposed model, the mobile sink node moves as per scheduling where it decides movement based on request counts.

The illustration of the considered network is presented in **Figure 2**.

Considering efficacy of expectations maximization based clustering model, we

consider it for initial clustering decision. A brief of the expectation maximization based centralized clustering mechanism is presented as follows:

#### 4.1. Expectation Maximization Algorithm Based CCP

Expectation maximization algorithm is a classical clustering approach, which considers that all the sensor nodes are distributed as per to GMM distribution. Mathematically

$$E(x) = \sum_{c=1}^C \sigma_c A(x | \sigma_c, \Sigma_c) \quad (1)$$

where  $C$  presents the total number of clusters, and  $\sigma_c$  signifies the mixing coefficient for  $c$ th cluster. The other parameters  $A(x | \sigma, \Sigma)$  can be mathematically derived as (2).

$$A(x | \sigma, \Sigma) = \frac{1}{(2\sigma) |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \sigma)^T \Sigma^{-1} (x - \sigma) \right\}, \quad (2)$$

where  $x$  states for the location vectors of the nodes. The other parameter  $\sigma_c$  presents the location vector of the CH of the  $c$ th cluster, while  $\Sigma_c$  represents the  $2 \times 2$  covariance matrix of the  $c$ th cluster.

In function, in the first stage EM model estimates the degree of dependence (DoD) of the individual node, called responsibility. DoD or responsibility factor (RF) depicts the dependence of a node on a cluster. In this way, the RF value of certain  $n$ th node on  $k$ th cluster can be estimated as follows:

$$\varphi_{nc} = \frac{\sigma_c a(x_n | \sigma_c, \Sigma_c)}{\sum_{j=1}^C \sigma_j a(x_n | \sigma_j, \Sigma_j)}. \quad (3)$$

Typically, the value of RF exists in the range of 0 and 1.

In the succeeding second stage of the EM based clustering, EM model estimates  $C$  weighted center of gravity (CoG) of a 2D-location vector for each node. This approach employs RF value as the weight of the node(s). Finally, in the third stage, the position of the CH are changed (or substituted) by the weighted CoG value as calculated in second stage. Thus, EM model estimates the log likelihood value by (4).

$$\mathbb{L} = \ln E(X | \sigma, \Sigma, \sigma) = \sum_{n=1}^N \ln \left\{ \sum_{c=1}^C \sigma_c A(x_n | \sigma_c, \Sigma_c) \right\} \quad (4)$$

The above mentioned process of the EM model continues iterating till reaches the convergence. In our proposed model the value of log likelihood is decreased uneventfully that makes termination of the EM model. Since, EM model updates CH information such as its position vector  $\sigma_c$ , and RF  $\varphi_{nc}$  of the individual nodes to the  $c$ th cluster, it results into gradual decrease in the sum of square (SoS) of the distances between individual node and cluster and thus eventually becomes optimal.

#### 4.2. Proposed Enhanced EM Based CCP Model

In this paper, we intend to enhance the generic EM based clustering model. As stated in this work we intend to develop a data gathering protocol for densely

distributed large scale geographical areas and therefore due to large number of sensors and its data volume, grouping of the nodes can be of paramount significance. Here, it should be noted that in our proposed model, “Group” states the set of nodes that can communication with each other. It states that a node can communicate only within its “Group” but can’t communicate with the nodes belonging to other “Group”. To perform data gathering from all connected nodes, it is inevitable to define the number of clusters more than the number of “Groups”.

In the proposed approach, initially the mobile sink allocates the CH randomly at a location. Applying the arbitrary position vector of the CHs, we have measured the communication distances  $D_{nc}$  between each node and the connected CHs and thus the mixing coefficient,  $\sigma$  and  $\Sigma$  are obtained. Once performing cluster initialization, the proposed model performs the selection of a group  $g$  using following Equation,

$$S_g = \frac{C_g}{N_g} \quad (5)$$

where,  $C_g$  presents the number of clusters in the group, while other parameter  $N_g$  the number of nodes in the group. Amongst the selected groups, the group with the highest  $S_g$ , we select all the connected sensor nodes from that group and updates the RF values  $\varphi_{nc}$  for those all connected node. Here,  $\varphi_{nc}$  signifies the extent to which the node  $n$  is connected to the cluster  $c$ . Applying updated RF ( $\varphi_{nc}$ ), the cluster heads CH ( $\sigma$ ), and the covariance matrix ( $\Sigma$ ), are re-measured, and thus the number of nodes belonging to the  $k$ th cluster is estimated using (6):

$$N_c = \sum_{x_i \in X} \varphi_{nc} \quad (6)$$

In our proposed model, the above mentioned computations continues iterating till the difference between the recently estimated  $E$  and the previous value of  $E$  becomes lower than the small number  $\gamma$ .

The overall proposed clustering model for data gathering is presented as follows

## 5. Enhanced EM (EEM) Based CCP Model for WSN Data Gathering

Once performing enhanced EM model for clustering, our proposed model executes the mobile sink node movement so as to patrol each CH and collect respective data. Here, it should be noted that the connected cluster nodes (CNs) or the sensor nodes transmits data to the mobile sink through CH. Here, it can be easily observed that with single mobile sink node transmission delay in WSN is the prime limitation. This significant delay can be stated as the waiting period in between the data generation to the data transmission from the sensor nodes. Since, the movement speed of the mobile sink is relatively slower as compared to the electrical communication between CNs and therefore there can be significantly long delay due to single mobile sink node. To deal with the delay prob-

lem, the reduction in the path trajectory (*i.e.*, patrolling distance) can be vital. It can enable mobile sink to reach CH early to collect its connected CN's data. In our model to achieve this, we have applied a heuristic approach named Traveling Salesman Problem (TSP) to estimate optimal (minimum distance) path for mobile sink node. Once reaching to the CH, the mobile sink collects data from CNs. Considering computational efficiency of the data collection algorithms, Directed Diffusion [36] is considered as one of the most efficient approach. With this motivation, we have applied its variant named "One Phase Pull approach [37] for data collection. In this approach, the mobile sink sends (data) transmission request message (TRM) to the CHs. Once receiving the TRM from cluster  $k$ , the mobile sink node re-transmits or re-broadcasts the data request and thus replies data to the adjacent node, which functions as the parent node in the data request tree of  $k$ . It is then followed by the relaying of the data from CNs to the mobile sink node. Here, it should be noted that one of the key novelties of the proposed data gathering scheme is dual constraints based transmission scheduling. As already stated that once moving to the CH, and getting TRM the mobile sink collects data; however to enable robust time efficient data gathering the relative distance between CN, and the CH or the mobile sink could not be addressed. Being greedy in nature WSN nodes intend to transmit data to the neighboring node to reach to the destination. Considering this fact, another scheduling measure has been incorporated in the proposed model where once getting TRM from CH and the mobile sink, the CNs estimates the relative distance and whoever comes first (lower intra-CH or intra-mobile sink) the CN relays its data to that. In this way, our proposed model avoids a significant waiting period and thus enables fast data gathering as expected for major real time applications.

On the other hand, to reduce the total energy required to transmit data, all connected nodes (CNs) transmit the sensed field data according to the RF value of the cluster. In our proposed data gathering approach, the RF value has been estimated on the basis of the network parameters  $\mu, \sigma$ , and  $\Sigma$ , as discussed in (3). These significant network parameters are appended to the data request message and are transmitted by the mobile sink. Once deploying CNs in the geographical region, the individual node performs exchange of its location vector  $x$ , with CNs belonging to the same groups. Since, the exchange of the key information such as the location vector  $x$  is performed only once after CNs deployment, it reduces the unwanted signaling overheads and thus reduces the energy consumption. However, it can't be stated as the optimal solution. When a sensor node belongs to a single cluster, it can transmit all data to the mobile sink directly. In case a node belongs to the multiple clusters, it transmits data as per the RF value of the individual cluster. For illustration, with  $\varphi_{n1} = 0.6$  and  $\varphi_{n2} = 0.4$ , in case the  $n$ th node receives TRM from mobile sink at the CH of the cluster 1, the CN transmits 60% of data to be transmitted. Similarly, in case a CN receives TRM sent from cluster 2, the CN transmits 40% of data to the mobile sink at the CH of the cluster 2. Thus, by transmitting data using the cluster adapted DD approach the total required energy can be minimized significantly. It can be found that the number of clusters does have the impact on the overall

network performance, especially on the network efficiency. The following section discusses the estimation of the optimal number of clusters so as to ensure optimal network performance.

### Optimal Cluster Count Estimation (OCCE) Model

In the previous sections of the presented manuscript, we emphasized on energy and delay efficient data gathering approach. However, considering an unavoidable questions whether the number of clusters does have impact on the efficiency, could not be addressed. In a number of existing works [13] authors have found that increase in the number of clusters may result into significant reduction in the energy consumption. Unfortunately most of the existing approaches don't consider energy consumption issues due to signaling process and TRM signaling. Considering these limitations, in this paper we have addressed such issues and derived a novel approach to estimate minimal number of clusters while maintaining higher throughput, minimal energy consumption and higher network efficiency.

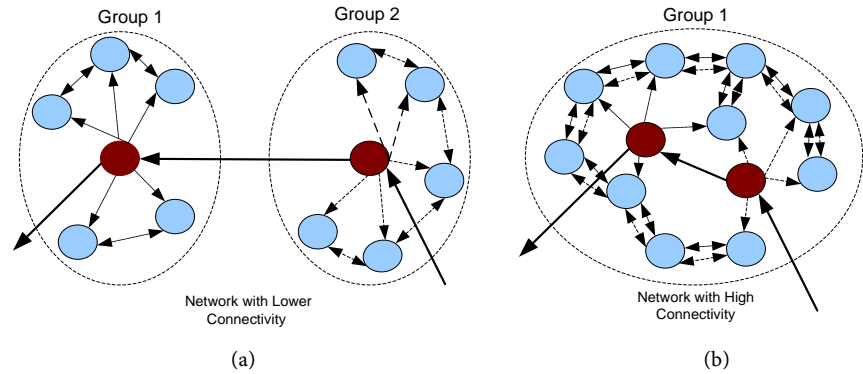
Being mobile in nature, in our proposed model, at first the inter-relationship between network connectivity and energy consumption has been examined. In order to investigate the correlation in between the node and sink connectivity, in our research we have derived a connectivity assessment model, in which connectivity is defined as a fraction of nodes which could communicate with each other. Mathematically connectivity is derived as follows (7).

$$CM = \frac{\sum_{g=1}^G N_g (N_g - 1)}{N(N-1)} \quad (7)$$

The connectivity matrix ( $CM$ ) possesses binary value (0 or 1), where if all nodes can communicate with each other,  $CM$  is updated with 1. On the contrary, if all the nodes are isolated then  $CM$  is assigned with 0. Whenever the mobile sink initiates process to estimate the optimal number of clusters needed to ensure energy efficient data gathering, it knows the CN's location across the network and thus it can estimate the connectivity parameter using Equation (7).

## 6. TRM Flooding

In practice, reaching the CH the mobile sink node transmits TRM to call up data transmission from CNs connected to that cluster. CNs receiving the request transmits its sensed data to the mobile sink and then broadcast TRM to its neighboring CNs. In our proposed model, the TRM continues till all connected CNs in the cluster has received it. In practice, there can be the possibility that a CN can get request twice or even more and in this situation at first it (*i.e.*, CN) transmits data and then broadcast TRM once after the first time of receiving the transmission request. Such uncontrolled transmission broadcast might introduce significantly higher energy exhaustion due to excessive (unwanted) redundant packet transmission. In such cases, minimizing such redundant transmission overheads or TRM is vital. In **Figure 3** the affect of TRM caused flooding can be visualized where it can be seen that with higher connectivity the issue of



**Figure 3.** Transmission request caused flooding in low as well as high connectivity WSN network.

flooding becomes prevalent. As depicted in the following figures **Figure 3(a)** **Figure 3(b)** there are two grouped in the network with lower connectivity, while there is a single group with the network with high connectivity. Here, it can be seen that there are 10 nodes distributed in the network region. In addition, the mobile sink traverses the two CHs, and then sends the TRM messages to the connected CNs. As depicted in **Figure 3(a)**, CNs can communicate only with each other in the same group (Group-1). The mobile sink transmits TRM to the individual node in cluster 1, and these connected CNs broadcast the data request message. In **Figure 3(b)**, where CNs can communicate with all nodes and hence the TRM transmitted at the cluster 1 get transferred to all the connected CNs. In addition, all CNs broadcast the data request message. Observing these two cases, it can be easily found that the requirements of the TRM for cluster 1 and cluster 2 is higher than that of cluster 1, as stated above.

Though the number of CNs and the clusters remain to be same, the severity of TRM flooding gets increased exponentially and severs, particularly with the network of higher connectivity **Figure 3(b)**. Furthermore, it becomes realizable that the overall TRM increases with increase in the number of clusters. Due to this problem, it becomes inevitable to estimate the optimal number of clusters while ensuring sufficient connectivity to have energy efficient data gathering. With this goal, in this paper we have derived a novel clustering count model that identifies the optimal number of clusters to perform efficient communication.

To estimate the optimal number of clusters in the distributed WSN, at first we have defined an objective function. In our proposed model we have considered the sum of required energy for TRM as well as data transmissions as the objective function  $Z(C)$ .  $Z(C)$  can be derived as the sum of energy consumption taken place due to the mobile sink patrolling (in one cycle). Mathematically,

$$Z(C) = L_{Req}S_{Req}(C) + L_{Dat}S_{Dat}(C), \quad (8)$$

where  $S_{Req}(K)$  and  $S_{Dat}(C)$  refers the SoS of transmission distance of data requests and data messages, respectively. Here, the parameters  $L_{Req}$  and  $L_{Dat}$  signify the size of the data and the size of the TRMs, correspondingly. Mathematically,  $S_{Dat}(C)$  is estimated using (9).



$$S_{Dat} = \sum_{n=1}^N \sum_{c=1}^C \sum_{h=1}^{H_{nc}} \phi_{nc} \cdot l_h^2 \quad (9)$$

where  $H_{nc}$  refers the total number of hops from  $n$ th node to the  $c$ th CH and  $l_h$  presents the communication distance of the individual hop. In the situation, when  $n$ th node is unable to communicate with the  $c$ th CH, in our model  $H_{nc}$  is set as 0, when the energy required is 0. Furthermore, the individual node re-transmits each TRM once with the highest transmission energy. Interestingly,  $S_{Dat}(C)$  used to be the decreasing function of  $C$  while TRM energy,  $S_{Dat}(C)$  is often increasing as a function of  $C$ , there exist a trade-off correlation in between the first and second component in the right component of (8). Now, realizing the functional condition that  $K$  should inevitably higher than the total number of groups  $G$ , the optimal number of clusters  $C_{optimal}$  can be obtained as (10)

$$C_{optimal} = \max\left(G, \arg \min_c (Z(C))\right) \quad (10)$$

To estimate the needed transmission power for TDMs, in our model we have considered the one group of nodes possessing  $N_g$  nodes and  $C_g$  CHs. TDMs is transmitted from each cluster and every connected CNs re-transmits it one time. In this way, the total needed energy for transmitting TDMs is obtained as:

$$S_{Req} = \sum_{g=1}^G C_g N_g M^2, \quad (11)$$

where  $M$  presents the highest radio range of the connected nodes (CNs) or the sensors. In case of non-imbalance of CHs location, the CNs belonging to the individual cluster is often same.

$$\frac{C_g}{N_g} = \frac{C}{N} \quad (12)$$

In case of more than 1 CN, the connectivity matrix ( $CM$ ) can be approximated as (13)

$$CM = \frac{\sum_{g=1}^G N_g (N_g - 1)}{N(N-1)} \doteq \frac{\sum_{g=1}^G N_g^2}{N^2} \quad (13)$$

Thus, applying above derived Equations (11-13),  $S_{Req}$  can be retrieved as (14).

$$S_{Req} = CN * M^2 C \quad (14)$$

This eventual model (14) states that the energy needed for TRMs transmissions (*i.e.*, data request transmission) is directly proportional to the connectivity. It reveals the fact that the number of clusters does affect the connectivity that eventually influences the energy consumption. Furthermore, the derived expression (14) signifies that the energy consumption increases as per the total number of clusters  $C$  that reflects that with lower clusters in the distributed geographical region can enable minimal energy exhaustion to perform data gathering. Thus, estimating the energy required for data transmission and gathering at the mobile sink (14), and TRM (9), the optimal clusters count can be obtained (10).

## 7. Inter Sensor Node-CH and Mobile Sink Distance Based Transmission Scheduling

In specific reference to the proposed energy efficient and delay sensitive data gathering using single mobile sink, in this work, we have incorporated a distance sensitive transmission scheduling model to perform fast data gathering at the mobile sink node. In the first case, as generic clustering based data gathering, the CNs transmits its data to the cluster head (CH) from which it belongs, which is then followed by the transmission of the data from CH to the mobile sink. On the contrary, the second approach of data transmission (*i.e.*, gathering at the mobile sink) exploits the relative distance between the CNs and the associated CH, and the nearest mobile sink. In case a node finds mobile sink nearer than the CH, the CN transmits its data directly to the mobile sink that not only reduces the computational overheads but also significant reduces delay, energy exhaustion and relaying cost etc. In this paper, the second case of implementation is stated to be the proposed system. Thus, applying this technique the delay sensitive and energy efficient routing model has been derived to achieve optimal data gathering in WSNs.

## 8. Results and Discussion

This section briefs about the simulation environment and the results obtained using proposed single mobile sink based data gathering protocol. The overall proposed work in this manuscript can be considered as a multiple constraints optimization model so as to achieve a novel solution for delay and energy efficient data gathering in WSN. At first considering the need of an efficient routing protocol, we emphasized on developing a robust single mobile sink based data gathering protocol. Further, exploiting the effectiveness of the centralized clustering paradigm (CCP) for a large scale network, which is more relevant for WSN infrastructures, we developed an enhanced expectation maximization (EM) based clustering model. No doubt, the enhanced EM model (say EEM) exhibited better results than the conventional K-Connid algorithm for data gathering. However, with the goal to further enhance the work, an additional constraint based on the distance between cluster nodes (CNs) and the CH as well as CNs and mobile sink was developed. In fact, the prime reason of introducing this distance sensitive data transmission scheduling was important so as to enable delay sensitive data gathering in WSNs. Unlike major existing approaches where generic clustering models have been incorporated, in this research work we developed an each-node's degree of dependence (DoD) based responsibility factor (RF) estimation that effectively enhances generic EM based clustering models. Furthermore, realizing the fact that the number of clusters formed in the network can have impact on the overall network performance for data gathering, in this work, novel optimal cluster count estimation (OCCE) model was developed. The prime objective of OCCE model was to examine the impact of the number of clusters on the energy exhaustion and the network-nodes (*i.e.*, CNs) lifetime. Thus, in this research effort has been made on enhancing both the op-

timal number of cluster selection, centralized clustering paradigm (CCP) enhancement, signaling overheads and associated energy exhaustion reduction, multi-constraints based data transmission scheduling model. Being a centralized clustering paradigm (CCP), the sensor nodes have the information about the mobile sink and the cluster heads (CHs), based on which data (transmission) routing is performed. We scheduled data gathering (from sensor nodes or CNs to the mobile sink) in two ways. In first approach, similar to the generic clustering based data gathering, CNs transmit data to the CH to which it belongs, which is further transmitted from CH to the mobile sink. On the contrary, the second approach of data transmission (*i.e.*, gathering at the mobile sink) exploits the relative distance between the CNs and the associated CH, and the nearest mobile sink. In case a node finds mobile sink nearer than the CH, the CN transmits its data directly to the mobile sink that not only reduces the computational overheads but also significant reduces delay, energy exhaustion and relaying cost etc. Thus, applying this technique the delay sensitive and energy efficient routing model has been derived to achieve optimal data gathering in WSNs. The simulation environment of the proposed model is presented in **Table 1**. The overall algorithms and simulation model is developed using Network Simulator (NS) version 2, commonly known as NS2. To present results in better perceptible manner, Matlab scripting has been applied for results plotting.

**Figure 4** presents the initial node deployment and grouping (here, we considered  $G = 2$ ). The purple color nodes signify one group while dark green color nodes are the CNs in other groups. Light green color node presents mobile sink node. Here, the individual node in each group can communicate with each other, while CNs from other group cannot communicate with CNs of other group.

The initial random CH selection for both the groups is depicted in **Figure 4**. Further, applying our proposed OCCE approach, the clusters have been obtained **Figure 5**.

Here, average minimum distance based CH selection has been done, in which each node estimates average distance in between itself and the other members belonging to the same cluster. The data transmission from CNs to the CH is presented in **Figure 6**. **Figure 6(a)** presents data gathering at initial instant,

**Table 1.** Simulation environment.

Parameters	Value
Total Number of Clusters (K)	4
Number of Node (N)	50
Communication Range (R)	200m
Field Length L	1000m
Clustering Algorithm	K-Conid, Enhanced Expectation Maximization, Proposed Clustering Algorithm
Number of Groups	2
Data Gathering	CNs→CH→Mobile Sink CNs→ Mobile Sink (Relative Distance Based Scheduling)

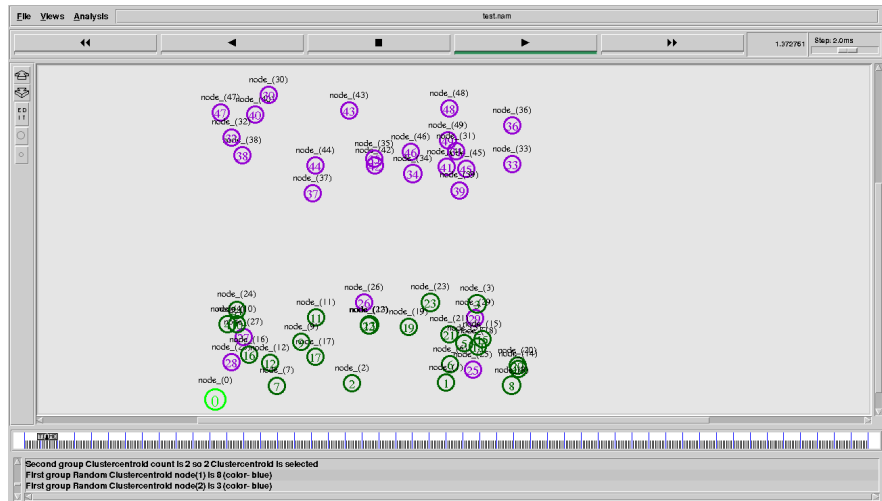


Figure 4. CNs distributions in two groups.

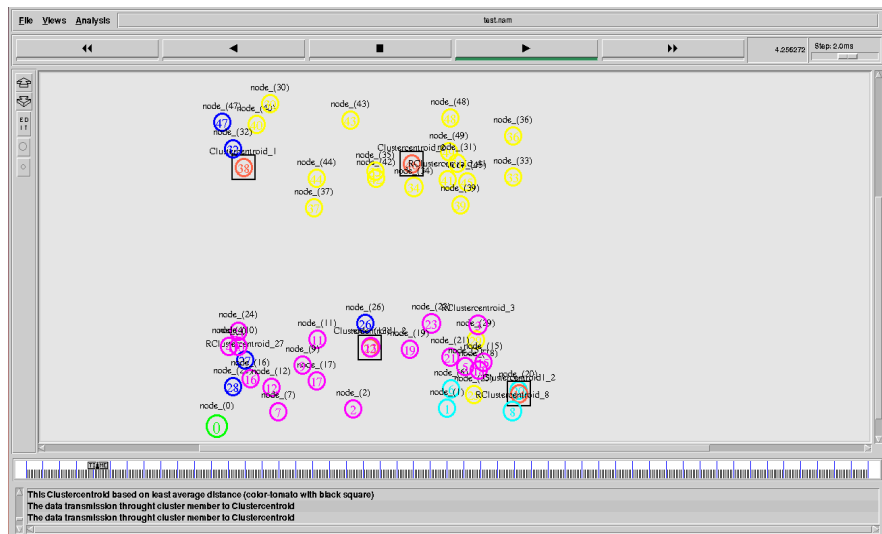
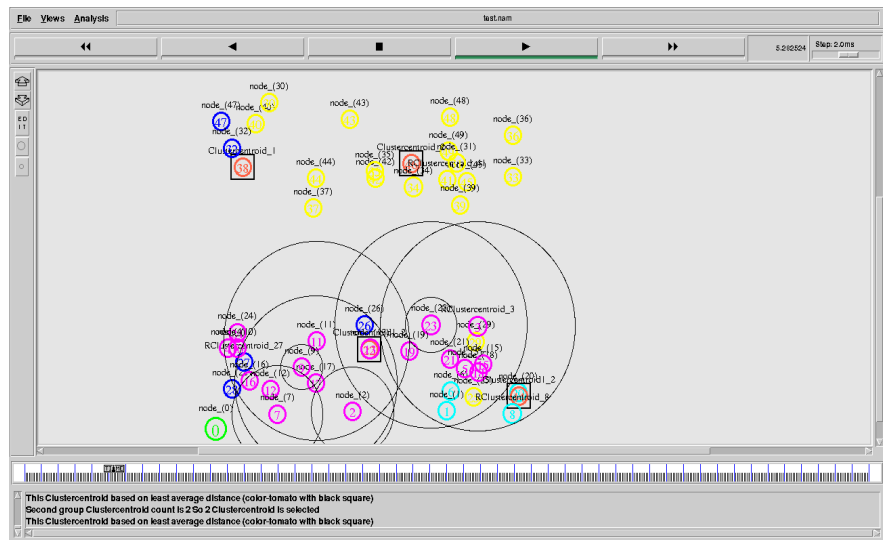


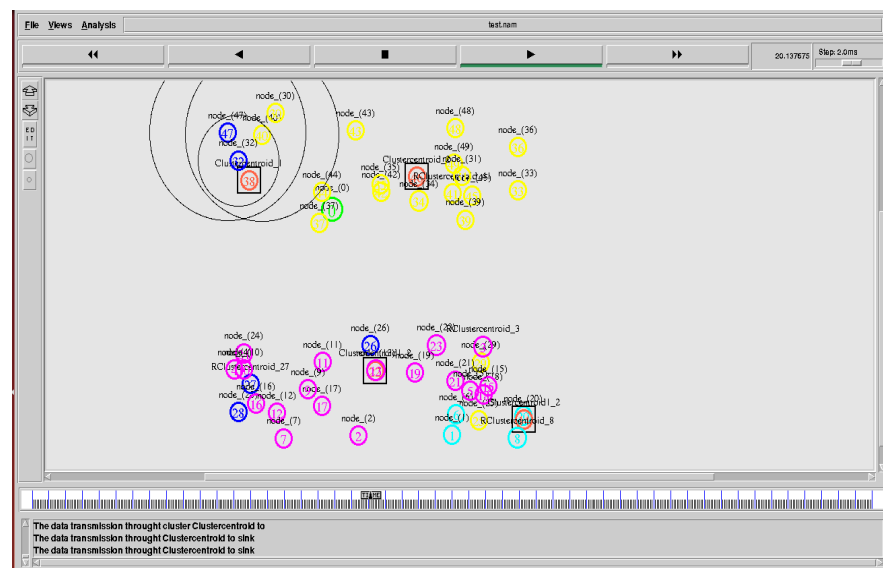
Figure 5. Cluster formation.

while being in mobility data gathering at certain ascending or proceeding instant is depicted in **Figure 6(b)**. Observing **Figure 6(a)** and **Figure 6(b)** the relative location of the mobile sink (NodeID-0, Color-Green) can be visualized. In further step as stated with the EEM based approach the CH forwards data to the mobile sink. The following snapshot presents the data gathering at the initial mobile sink position, which varies as per random movement pattern.

In the experiment process, performing optimal clustering and introducing mobile sink based data gathering mechanism; the energy consumption for the data transmission ( $E_{Dat}$ ) from CNs to the mobile sink has been done. Here,  $E_{Dat}$  signifies the amount of energy required to perform data transmission from sensor nodes or the connected nodes (CNs) to the mobile sink. Here, in case the position of CH is far away (beyond the communication range of a CN) and it can't make connection with CH, the  $E_{Dat}$  value is updated with 0. In fact, this situation reflects failure of clustering and doesn't mean the energy conservation.



(a)



(b)

**Figure 6.** (a). Data gathering at initial instant; (b). Data gathering during mobility.

The research simulation with the generic expectation maximization based clustering which doesn't consider any connectivity issue undergo degraded performance as reflected through results presented below. Here, K-Conid algorithm based clustering which applies generic EM based clustering for data transmission is considered as a reference model. However, as our proposed model, we have at first enhanced EM based clustering (here we say, EEM) to perform data gathering. In addition, based on transmission strategy (as already discussed that to enable fast data gathering we have applied CH and mobile sink location information to make transmission scheduling), we have examined our proposed model in two approach. In first, with Enhanced Expectation Maximization (EEM) based clustering followed by optimal cluster count estimation (OCCE) and transmission in "CNs→CH→Mobile Sink" manner. In the second contribution,

with intend to enable fast data gathering at sink, we have considered relative distance amongst CNs, CH and mobile sink, where transmission is scheduled as “CNs→Mobile Sink” (Relative distance based scheduling). Here, in result discussion the second data gathering case (CNs→ Mobile Sink) is referred as the proposed system.

The following figures presents the results obtained from NS2 simulation, which are encoded in Matlab scripting and are presented so as to make more perceptible.

**Figure 7** presents the performance of the proposed data gathering algorithms in terms of the packet delivery ratio (PDR). Here, it can observed that the proposed EEM model which is the enhanced version of the generic EM clustering based data gathering performs satisfactory. Interestingly, the data gathering model with (CNs→ Mobile Sink) outperforms “CNs→CH→Mobile Sink” transmission model. Though, these two approaches employ same clustering model and enhanced cluster selection model for delay energy efficient transmission; however the proposed model with direct data delivery (*i.e.*, CNs→Mobile Sink) performs batter than via CH node in the cluster. This can be due to additional losses incurred due to receiving and transmission at the CH node. It should be noted that in this paper, the relative distance between CN and the cluster head (CH) of that cluster and mobile sink was used to decide whether the CN should transmit its data through CH or directly to the mobile sink. In case the distance between CN and mobile sink is less than the distance between CN and CH and the mobile sink comes under the communication range of the CN, then the CN prefers transmitting its data directly to the mobile sink. In addition, the result **Figure 7** reveals that with increase in the number of nodes in each cluster, the packet delivery ratio (PDR) decreases. This is because of the increase in data traffic load over CH. Undeniably, WSN nodes being greedy in nature can cause significantly higher contention at the CH and thus the rate of data drop can increase due to collision and contention. Such evidences could be seen in **Figure 6**, where black-color dots present data drop due to congestion at that cluster head (CH). It signifies that the selection of the optimal number of cluster heads (CHs) is must to ensure higher throughput and reliable data gathering process. In reference to the congestion caused (due to increase in the number of nodes per cluster), the retransmission caused delay can be easily visualized in **Figure 8**. However, it signifies that the proposed data gathering scheme (EMM based CCP and CNs→ Mobile Sink data gathering) exhibits better as compared to the EMM based data gathering.

This is the fact that the inter-CN/CH/mobile sink distance approach was applied only to enable fast data gathering by avoiding unwanted traversal from sensor node or cluster node (CN) to the CH and then from CH to mobile sink. Transmitting data from CN to the mobile sink can't not only reduce the probability of packet drop but also alleviates unwanted traversal and hence reduces delay in data gathering. **Figure 8** exhibits that the proposed transmission model can perform fast data gathering, which is of paramount significance to ensure

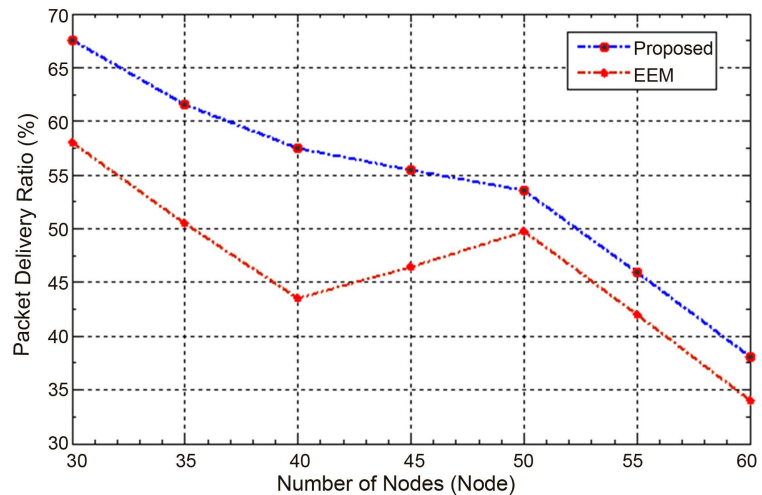


Figure 7. Packet delivery ratio.

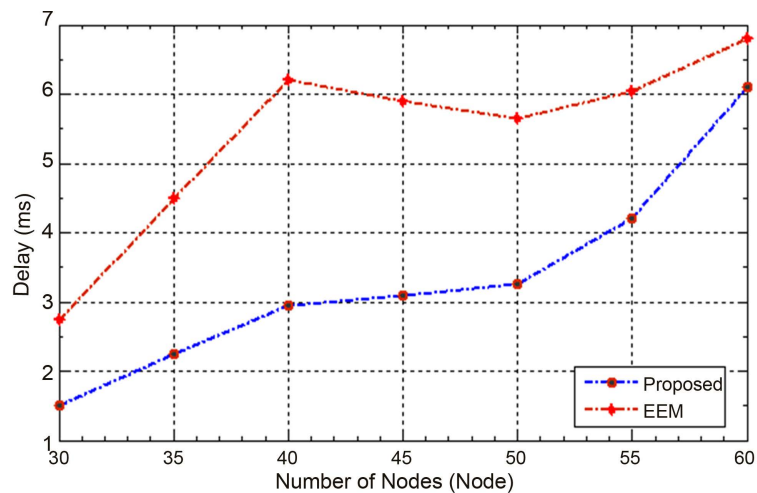
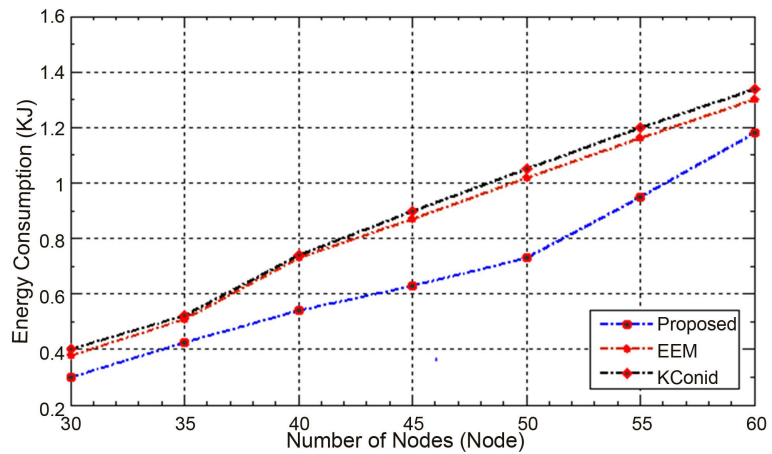


Figure 8. Transmission or data gathering delay.

delay sensitive communication over WSNs.

Considering energy efficiency of the proposed data gathering model, it can be seen **Figure 9** that the proposed data gathering protocol outperforms others, including EEM and K-Conid based approaches. No doubt, the reduction in data drop, higher PDR, lower retransmission probability, lower energy exhaustion in clustering, CH selection process, and TRM enable our proposed routing scheme to exhibit minimal energy consumption.

For any routing protocol, its computational efficiency plays a vital role to enable it for real world applications. In this paper, we have examined the performance of the proposed routing protocol in terms of efficiency. In addition, in this paper an additional factor was considered called efficiency. Considering discussions in previous section, it is visible that those clustering approaches that don't consider connectivity undergo significant failure causing data drop, retransmission, delay and energy consumption. Considering these facts, in this paper a metric called "Efficiency" was introduced which was calculated as (15):



**Figure 9.** Energy consumption by different techniques.

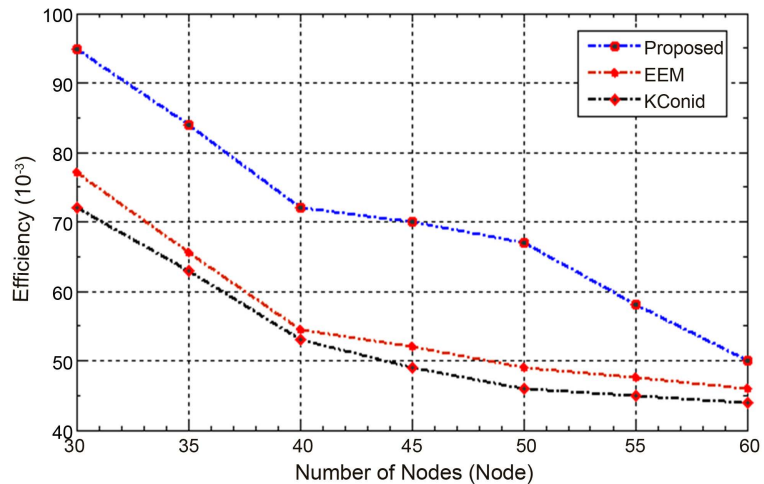
$$\text{Efficiency} = \frac{\text{Number of connected node}}{S_{Dat}} \quad (15)$$

The performance assessment of the proposed data gathering scheme with existing distributed clustering scheme K-Conid and our proposed CCP based clustering and data gathering, it can be found **Figure 10** that the proposed data gathering approach exhibits higher efficiency than the other approaches.

## 9. Conclusion

The exponential rise in the communication system demands has motivated academia-industries to develop certain low-cost, energy efficiency and QoS oriented communication systems. Wireless Sensor Network (WSN) has always been the dominating technology serving an array of solutions including civil, defense and industrial monitoring, control and decision purposes. However, enabling QoS delivery, delay sensitive and energy efficient data gathering has been the key domain for researchers. In this paper, the significance of single mobile sink was exploited to perform data gathering in WSN. Considering robustness of centralized clustering paradigm or approach, in this paper an Enhanced Expectation-Maximization model (EEM) based clustering model was developed. Observing the fact that the number of clusters impacts on the energy consumption, a novel optimal cluster count estimation (OCCE) model was developed that intended to ensure minimal energy exhaustion particularly caused due to signaling overheads (data transmission request and re-broadcasting across nodes). Unlike existing approaches, in this paper a relative inter-node distance (distance between sensor nodes and cluster heads, and the distance between sensor nodes and mobile sink) based transmission scheduling model was developed, which enabled the final proposed data gathering protocol to exhibit higher throughput, minimal delay and energy consumption, and higher efficiency than other distributed clustering based approaches and even EEM based clustering with generic transmission mechanism. Since, the movement pattern and decision of the mobile sink towards cluster does have impact on reliable and delay sensitive data





**Figure 10.** Network communication efficiency.

gathering. In this paper, the sink movement was considered as travelling salesman problem. In future network condition aware scheduling can be done to reduce unwanted patrolling across network or to reduce data drop at a cluster due to unavailability of mobile sink.

## References

- [1] Hamida, E.B. and Chelius, G. (2008) Strategies for Data Dissemination to Mobile sinks in Wireless Sensor Networks. *IEEE Wireless Communications*, **15**, 31-37. <https://doi.org/10.1109/MWC.2008.4749745>
- [2] Ye, F., Luo, H., Cheng, J., Lu, S. and Zhang, L. (2002) A Two-Tier Data Dissemination Model for Large-Scale Wireless Sensor Networks. *Proceedings of the 8th Annual International Conference on Mobile Computing and Networking*, Atlanta, Georgia, 23-28 September 2002, 148-159. <https://doi.org/10.1145/570645.570664>
- [3] Sha, Z., Lu, J., Li, X. and Wu, M. (2010) An Anti-Detection Moving Strategy for Mobile Sink. *IEEE Proceeding of Global Telecommunications Conference*, Miami, FL, 6-10 December 2010, 1-5. <https://doi.org/10.1109/glocom.2010.5684043>
- [4] Cheng, L., Chen, Y., Chen, C. and Ma, J. (2009) Query-Based Data Collection in Wireless Sensor Networks with Mobile Sinks. *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing. Connecting the World Wirelessly*, Leipzig, Germany, 21-24 June 2009, 1157-1162. <https://doi.org/10.1145/1582379.1582633>
- [5] Kinalis, A., Nikolettseas, S., Patroumpa, D. and Rolim, J. (2014) Biased Sink Mobility with Adaptive Stop Times for Low Latency Data Collection in Sensor Networks. *Information Fusion*, **15**, 56-63.
- [6] Park, T., Kim, D., Jang, S., Yoo, S. and Lee, Y. (2009) Energy Efficient and Seamless Data Collection with Mobile Sinks in Massive Sensor Networks. *IEEE International Symposium on Parallel & Distributed Processing*, Rome, 23-29 May 2009, 1-8. <https://doi.org/10.1109/ipdps.2009.5161183>
- [7] Sun, Z., Zheng, Z., Xu, S. and Chen, S. (2010) Energy-Aware Data Gathering Protocol for Wireless Sensor Networks. *3rd IEEE International Conference on Computer Science and Information Technology*, Chengdu, 9-11 July 2010, 103-107.
- [8] Boukerche, A. and Nelem Pazzi, R.W. (2007) Lightweight Mobile Data Gathering

Strategy for Wireless Sensor Networks. *9th IFIP International Conference on Mobile Wireless Communications Networks*, Cork, 19-21 September 2007, 151-155.

<https://doi.org/10.1109/ICMWCN.2007.4668199>

- [9] Taifeh Aliverdi, M. and Pakravan, M.R. (2007) A Network Information Gathering (NIG) Protocol and Energy-Aware Load Balancing Algorithm for Wireless Sensor Network. *IEEE International Conference on Telecommunications and Malaysia International Conference on Communications*, Penang, 14-17 May 2007, 407-412.
- [10] Gherbi, C., Aliouat, Z. and Benmohammed, M. (2015) Distributed Energy Efficient Adaptive Clustering Protocol with Data Gathering for Large Scale Wireless Sensor Networks. *12th International Symposium on Programming and Systems*, Algiers, 28-30 April 2015, 1-7. <https://doi.org/10.1109/isps.2015.7244966>
- [11] Shah, M.B., Tamhankar, P.S., Merchant, S.N. and Desai, U.B. (2011) A Realistic Weighted Clustering Algorithm for Data Gathering in Single Hop Cell Phone Based Sensor Network. *IEEE GLOBECOM Workshops*, Houston, TX, 5-9 December 2011, 1253-1257. <https://doi.org/10.1109/glocomw.2011.6162382>
- [12] Almi'ani, K., Viglas, A. and Libman, L. (2010) Energy-Efficient Data Gathering with Tour Length-Constrained Mobile Elements in Wireless Sensor Networks. *35th Conference on Local Computer Networks*, Denver, CO, 10-14 October 2010, 582-589. <https://doi.org/10.1109/lcn.2010.5735777>
- [13] Takaishi, D., Nishiyama, H., Kato, N. and Miura, R. (2014) Toward Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks. *IEEE Transactions on Emerging Topics in Computing*, **2**, 388-397. <https://doi.org/10.1109/TETC.2014.2318177>
- [14] Lee, D., Park, S., Lee, E., Choi, Y. and Kim, S.-H. (2007) Continuous Data Dissemination Protocol Supporting Mobile Sinks with a Sink Location Manager. *Proceedings of the Asia-Pacific Conference on Communications*, Bangkok, 18-20 October 2007, 299-302.
- [15] Nesamony, S., Vairamuthu, M.K. and Orłowska, M.E. (2007) On the Traversals of Multiple Mobile Sinks in Sensor Networks. *Proceedings of the IEEE International Conference on Telecommunications and Malaysia International Conference on Communications*, Penang, 14-17 May 2007, 432-437. <https://doi.org/10.1109/ICTMICC.2007.4448675>
- [16] Song, L. and Hatzinakos, D. (2005) Dense Wireless Sensor Networks with Mobile Sinks. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP05)*, **3**, 677-680.
- [17] Wang, Z.M., Basagni, S., Melachrinoudis, E. and Petrioli, C. (2005) Exploiting Sink Mobility for Maximizing Sensor Networks Lifetime. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 6 January 2005, 287a. <https://doi.org/10.1109/hicss.2005.259>
- [18] Traynor, P., Shin, J.S., Madan, B., Phoha, S. and La Porta, T. (2006) Efficient Group Mobility for Heterogeneous Sensor Networks. *Proceedings of the IEEE 64th Vehicular Technology Conference*, Montreal, 25-28 September 2006, 1-5. <https://doi.org/10.1109/VTCF.2006.545>
- [19] Qiu, Q. and Kamal, A.E. (2005) Coverage and Connectivity Control of Wireless Sensor Networks under Mobility. *Workshop on High Performance Switching and Routing*, 12-14 May 2005, 177-181.
- [20] Jea, D., Somasundara, A. and Srivastava, M. (2005) Multiple Controlled Mobile Elements (Data Mules) for Data Collection in Sensor Networks. In: Prasanna, V.K., Iyengar, S.S., Spirakis, P.G. and Welsh, M., Eds., *Distributed Computing in Sensor Systems. DCOSS 2005. Lecture Notes in Computer Science*, Vol. 3560, Springer,

- Berlin, Heidelberg, 244-257. [https://doi.org/10.1007/11502593\\_20](https://doi.org/10.1007/11502593_20)
- [21] Gandham, S.R., Dawande, M., Prakash, R. and Venkatesan, S. (2003) Energy Efficient Schemes for Wireless Sensor Networks with Multiple Mobile Base Stations. *Proceedings of the IEEE Global Telecommunications Conference*, **1**, 377-381. <https://doi.org/10.1109/GLOCOM.2003.1258265>
- [22] Madhumathy, P. and Sivakumar, D. (2014) Enabling Energy Efficient Sensory Data Collection Using Multiple Mobile Sink. *China Communications*, **11**, 29-37. <https://doi.org/10.1109/CC.2014.6969791>
- [23] Zhu, C., Wu, S., Han, G., Shu, L. and Wu, H. (2015) A Tree-Cluster-Based Data-Gathering Algorithm for Industrial WSNs with a Mobile Sink. *IEEE Access*, **3**, 381-396. <https://doi.org/10.1109/ACCESS.2015.2424452>
- [24] Wang, W., Srinivasan, V. and Chua, K.C. (2005) Using Mobile Relays to Prolong the Lifetime of Wireless Sensor Networks. *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking*, Cologne, Germany, 28 August-2 September 2005, 270-283. <https://doi.org/10.1145/1080829.1080858>
- [25] Guo, S., Wang, C. and Yang, Y. (2014) Joint Mobile Data Gathering and Energy Provisioning in Wireless Rechargeable Sensor Networks. *IEEE Transactions on Mobile Computing*, **13**, 2836-2852. <https://doi.org/10.1109/TMC.2014.2307332>
- [26] Ma, M., Yang, Y. and Zhao, M. (2013) Tour Planning for Mobile Data-Gathering Mechanisms in Wireless Sensor Networks. *IEEE Transactions on Vehicular Technology*, **62**, 1472-1483. <https://doi.org/10.1109/TVT.2012.2229309>
- [27] Liu, D., Zhang, K. and Ding, J. (2013) Energy-Efficient Transmission Scheme for Mobile Data Gathering in Wireless Sensor Networks. *China Communications*, **10**, 114-123.
- [28] Cheng, C.F. and Lee, H. (2016) Data Gathering in Wireless Sensor Networks with Uncontrolled Sink Mobility. 2016 *IEEE 83rd Vehicular Technology Conference*, Nanjing, 15-18 May 2016, 1-5. <https://doi.org/10.1109/vtcSpring.2016.7504183>
- [29] Huang, J. and Liu, D. (2013) A High-Reliability Data Gathering Protocol Based on Mobile Sinks for Wireless Sensor Networks. *22nd Wireless and Optical Communication Conference*, Chongqing, 16-18 May 2013, 304-308. <https://doi.org/10.1109/wocc.2013.6676327>
- [30] Shah, R.C., Roy, S., Jain, S. and Brunette, W. (2003) Data MULEs: Modeling and Analysis of a Three-Tier Architecture for Sparse Sensor Networks. *Ad Hoc Networks*, **1**, 215-233.
- [31] Heinzelman, W., Chandrakasan, A. and Balakrishnan, H. (2000) Energy Efficient Communication Protocol for Wireless Microsensor Networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, **2**, 10. <https://doi.org/10.1109/hicss.2000.926982>
- [32] Youssef, M., Youssef, A. and Younis, M. (2009) Overlapping Multihop Clustering for Wireless Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, **20**, 1844-1856. <https://doi.org/10.1109/TPDS.2009.32>
- [33] Mai, K.T. and Choo, H. (2008) Connectivity-Based Clustering Scheme for Mobile Ad Hoc Networks. *IEEE International Conference on Research, Innovation and Vision for the Future*, Ho Chi Minh City, 13-17 July 2008, 191-197.
- [34] Lindsey, S. and Raghavendra, C. (2002) PEGASIS: Power-Efficient Gathering in Sensor Information Systems. *IEEE Aerospace Conference*, **3**, 3-1125-3-1130. <https://doi.org/10.1109/aero.2002.1035242>
- [35] Nakayama, H., Ansari, N., Jamalipour, A. and Kato, N. (2007) Fault-Resilient Sensing in Wireless Sensor Networks. *Computer Communications*, **30**, 2375-2384.

- [36] Intanagonwiwat, C., Govindan, R. and Estrin, D. (2000) Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, Boston, MA, 6-11 August 2000, 56-67. <https://doi.org/10.1145/345910.345920>
- [37] Chen, M., Kwon, T. and Choi, Y. (2006) Energy-Efficient Differentiated Directed Diffusion (EDDD) in Wireless Sensor Networks. *Computer Communications*, **29**, 231-245.



Scientific Research Publishing

**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact [wsn@scirp.org](mailto:wsn@scirp.org)

