Enhanced Single Mobile Sink Traversibility Based Delay Sensitive Data Gathering in Large Scale Wireless Sensor Networks

# Shivkumar S.Jawaligi

Electronics & Communication Engineering PDA College of Engineering

Kalaburagi, India shiv.jawaligi@gmail.com

# G. S. Biradar

Electronics & Communication Engineering PDA College of Engineering

Kalaburagi, India gsbiradar@gmail.com

***Abstract—In this paper, a delay sensitive, energy efficient and quality of service (QoS) oriented WSN data gathering protocol has been developed using single mobile sink node. We exploit an enhanced centralized clustering model using expectation-maximization (EEM) concept to enable reliable and energy efficient clustering based data transmission. To reduce energy exhaustion and signaling overheads due to more number of clusters, we derived an optimal cluster count model. To enable time-efficient WSN data gathering the use of inter-node distance*** (𝑑i𝑠𝑡(𝐶𝑁𝑠𝐶𝐻𝑀𝑜𝑏i𝑙e 𝑆i𝑛k)𝑜𝑟 𝐶𝑁𝑠 𝑀𝑜𝑏i𝑙e 𝑆i𝑛k) ***based dual transmission strategy has exhibited optimal data gathering. The overall 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, Centralized Clustering, Expectation- Maximization.***

incorporates multiple nodes distributed across network region where each node senses the events data and transmits to the base station (BS) through neighboring sensor nodes. To ensure optimal QoS delivery, the efficiency of data gathering scheme and best forwarding node or path selection play vital role [1-4]. However, in traditional WSN communication paradigm, data gathering occurs through multi-node traversal mechanism, where sensor node transmits data (towards sink) through its neighboring node fulfilling certain node-criterion as illustrated in Fig. 1.

1. INTRODUCTION

In last few years, huge demand has been witnessed for an efficient wireless communication system to serve civil communication purposes, surveillance systems, administrative utilities, critical data communication across interfaced operating units, defense systems, business communications, and industrial communication, etc. On the other hand, the exponential rise in wireless communication technologies and associated demands has given rise to a new research horizon to develop certain more efficient and productive communication system to meet major demands, as mentioned above. “Internet of Things” typically stated as IoT technology has gained significant attention globally to enable more efficient and productive communication system. Facilitating optimal Quality of Service (QoS) inevitably demands low cost solution with higher bandwidth utilization, minimum end-to-end delay, minimal energy consumption, reliable communication etc. The robustness of the wireless sensor network (WSN) enables it to be one of the most sought networking technologies to serve major wireless communication demands and IoT service provision. WSN

Fig. 1. Traditional Wireless Sensors Network.

This generic process not only introduces high end-to-end delay and computational overheads, but also causes energy exhaustion and thus limiting overall network-efficiency [5]. No doubt, the efficacy of the multi-hop transmission has been acknowledged in many literatures [1-2] [6]; however delay incurred has also been indicated in major existing approaches. To alleviate the major existing issues such as multi-hop traversal based data transmission, delay, data drop

**979-8-3315-4317-4/25/$31.00 ©2025 IEEE**

probability etc recently mobile sink based data gathering technique has been suggested [1-2][6-9][10]. The implementation of the mobile sink based data gathering assures higher throughput, minimal data drop and hence retransmission probability and more importantly timely data delivery [10, 11] [12-19]. As adversaries, introducing mobility in WSN causes exceedingly higher topological changes that give rise to congestion probability, network contention, data drop or data lifetime failure likelihood etc. The efficacy of multiple mobile sink based data gathering has been applauded [20]; however the computational overheads incurred often remains a question for these approaches.

To deal with such issues, the use of single mobile sink node has been found a better and potential alternative [21- 24]. In a number of researches, for a large scale network clustering based routing has been found a better alternative [21][25][26] to ensure reliable data transmission with higher resource utilization. To enable QoS delivery in large scale network, mobile sink based approaches have been found effective [27], where to further enhance network efficiency, splitting network into multiple grids have been suggested. However, with single mobile sink, traversing different grids and collecting timely data from each grid has been found intricate. To deal with this deciding optimal path selection, data and deadline sensitive traversal strategies is must. Exploring in depth, it can be found that the majority of the distributed schemes like LEACH apply the limitation of the node’s communication range in the network that often fails in mobile sink based data gathering [26]. Recently developed distributed clustering scheme named K-hop Overlapping Clustering Algorithm (KOCA) [28] and k-hop connectivity ID (k-CONID) [29] have been found effective for clustering based WSN data gathering, in which KOCA focuses on multiple overlapping clusters, and applies probabilistic model based cluster head (CH) selection for data transmission Distinctly, in k-CONID the nodes exhibit their IDs exchange with each other in random fashion, and the node having minimum ID within k-hop is chosen to be the CH, through which data transmission takes place. With the goal to develop a more efficient and effective routing scheme this paper introduces a centralized clustering model (CCM) based data gathering model. Realizing the efficiency of CCM approaches such as Power-efficient data gathering in Sensor Information Systems (PEGASIS) [30], this paper applies travelling salesman problem (TSP) based mobile sink movement which enables minimal node traversal and hence negligible delay than the existing approach [29]. To enable efficient routing and associated data gathering, authors [31] performed clustering using K-Means algorithm, which was then followed by distance based traversibility scheduling for single mobile sink based data gathering. However node movement irrespective of the node’s location limits its effectiveness to enable delay sensitive data gathering [26].

This paper exploits the effectiveness of the clustering technique to perform grouping of the nodes (i.e., clustering) based on their geographical locations across the large scale network. The proposed work applies single mobile sink based data gathering that avoids any forwarding nodes or

multi-hop transmission and thus minimizes channel load, contention and computational complexities. Unlike multiple mobile sink nodes based data gathering, this paper applies single mobile sink to perform data gathering from different connected (cluster) nodes. It as a result reduces huge computational overheads, contention, unwanted traversal caused delay and energy exhaustion. In this paper an enhanced data transmission model has been developed that at first employs robust heuristic approach to decide best node mobility path or traversal path so as to enable fast sink movement to collect data from the requesting CH. Being an enhanced expectation maximization (EMM) based clustering model, our approach applied a dual transmission strategy that performs data collection from CH as well as directly from the connected sensor nodes (CNs) in the cluster. It significantly strengthens the fast data gathering.

The other sections of the presented research are divided as follows: Section II presents the brief of the proposed contribution, which is followed by system model and algorithmic implementation (discussion) in Section III. The experimental results and its discussion are presented in Section VII. Overall conclusion and future scopes are presented in Section VIII. References used are given at the end of the manuscript.

1. PROPOSED MODEL

In this research work, the emphasis has been made on developing a novel EEM based clustering model to perform delay sensitive and QoS oriented data gathering in large scale WSNs using single mobile sink node. In this work, the emphasis has been made on the following key targets:

1. Exploiting the significance of single mobile sink node to perform QoS delivery, delay sensitive, and energy efficient data gathering in WSNs,
2. Investigating the affect of control messages on the energy consumption and its reduction strategy,
3. Identifying the optimal number of clusters so as to maintain minimum energy exhaustion and computational overheads,
4. To derive each-node responsiveness factor for CH selection and dynamic transmission path decision so as to enable fast data gathering,
5. To derive a dual transmission model strategies for fast data gathering in large scale WSN.
6. Assess the performance of the proposed routing scheme and compare with other available CCM based WSN data gathering protocols.

A brief of the proposed WSN data gathering model using single mobile sink node is given as follows:

*A. Centralized Clustering Model (CCP) Based Data gathering*

To develop an efficient data gathering model using single mobile sink, the main issue is to minimize energy consumption, which can be achieved by means of estimating optimal location where data gathering has to be done. In this work, a hypothesis that the required transmission energy by a node is always proportional to the square of the transmission

distance has been considered. In other words, we focus on reducing the sum of square (SoS) of transmission distance so as to reduce energy consumption in WSN data gathering. Expectation Maximization (EM) that applied iterative mathematical calculations to perform clustering is applied in our model. Realizing the robustness of the EM approach that significantly minimizes the SoS of the inter-node distance and the CH, in this work EM has been applied. Practically there exist limitations, such as a node can perform transmission only within its communication range and not every sensor nodes can communicate with each other or even CH. In such conditions, the nodes which are unable to communication with CH require communicating in multi- hop approach. For multi-hop communication, the communication range states the SoS between the sensor nodes occurring in the path towards BS. It reveals that the communication range is always different from the direct distance between the sensor node and the BS (Fig. 2). However, EM model often focuses on minimizing the SoS of the direct distance, rather minimizing the communication range. In this situation, applying EM model is vital to perform data gathering in WSN where each node possesses certain fixed communication range. The following section discusses the overall proposed routing scheme for data gathering is using single mobile sink node.

1. SYSTEM MODEL

As stated, the proposed model comprises single mobile sink node and multiple sensor nodes placed across the network region. Here, we assume that the individual node is aware about its location and mobile sink too possesses the information about each nodes distributed across the network. The individual node has the communication range 𝑅 within which it can communicate successfully. Being mobile sink the mobile node patrols CHs so as to avoid multipath transmission and associated data drop probability, which as a result reduces energy consumption significantly. The individual node possesses a defined sized buffer so as to store received network information till mobile sink reaches to CH to collect it. In this way, the information gathered at each

belonging to the different groups can’t communicate with each other. The variable 𝐺 signifies the total number of groups in the network region. The other variables 𝑁g and 𝐶g present the total sensor nodes and the number of clusters in the 𝑔th group, respectively. The total number of groups is estimated by means of the key information such as node location and their communication range 𝑅. In the proposed data gathering protocol, the mobile sink traverses on the basis of the request counts from each CH.

 

Fig. 2. An illustration of the considered WSN network

 2

2

As stated above, in this research EM based centralized clustering model (CCM) is applied to perform data gathering:

1. *Expectation Maximization (EMM) Based CCM*

Expectation maximization is a generic clustering model that considers sensor nodes distributed accordingly to Gaussian Mixture Model (GMM). Mathematically

node is transferred to the mobile sink. Interestingly, we have incorporated a factor called “Node-Responsiveness (NR)” that enables the dual transmission strategy decision process. In our proposed routing model, NR decides whether it should

𝐶

𝐸(x) = ∑ 𝜎𝑐𝐴(𝑥|𝜎𝑐, Σ𝑐)

𝑐=1

(1)

transmit data through CH or directly to the mobile sink which is nearer than the CH. Here, it should be noted that the prime objective of NR is to enable fast data gathering in

Where the parameters 𝐶 and 𝜎𝑐 refer the total number of

clusters the mixing coefficient for 𝑐 th cluster. The parameter

𝐴(𝑥|𝜎, Z) is derived as (2).

WSN.

In our model, WSN is considered to be densely distributed characterizing urban communication environment. Considering large scale network, we divide

𝐴(𝑥|𝜎, Z) = 1 e𝑥𝑝 {− 1 (𝑥 − 𝜎)𝑇Z−1(𝑥 −

2𝜎 |Z| 2

( ) 1/2

𝜎)}*,*

(2)

overall network region into sub-network regions. In Fig. 2, circles refer the 𝑁 sensor nodes distributed throughput network dimension 𝐿 × 𝐿 , where 𝐿 signifies geographical network parameter. The solid color filled circle 𝐶 refers the CHs supposed to be visited by the mobile sink to collect its data. The solid-fill region is the group of the nodes, while dotted circle presents the cluster. We state the term “group” as the set of 𝑛 nodes able to perform communication with each other. Due to inter-node distance factor, the nodes

where, the parameter x and σc signify the location vectors for nodes and CHs respectively. Σc refers 2×2 covariance matrix of the c th cluster.

Functionally, at first EM approach calculates the degree of dependence (DoD) of each node, also called “Node Responsibility (NR)”. Here, NR presents the node dependency on a cluster. Thus, the NR value of 𝑛 th node on kth cluster is obtained by (3).

𝑁𝑅 = 𝜎𝑐𝑎(𝑥𝑛|𝜎𝐶, Σ𝐶) . (3)

∑

𝑁𝑐 = ∑ 𝑁𝑅

(6)

𝑐 j=1

𝜎j𝑎 (𝑥𝑛|𝜎j, Σj)

𝑥𝑛∈𝑋

where NR exists in the range of 0 to 1.

In the next phase of the EM clustering, we calculate C weighted center of gravity (CoG) of the two dimensional location vector of the individual node. This approach considers NR value as the weight of each node. Finally, in third stage, the location vectors of the CHs are replaced by the weighted CoG value. Thus, EM calculates the log likelihood value using (4).

The above processes continue iterating until the difference between the recently estimated E and the previous value of E becomes lower than the small number γ . The overall proposed clustering model is presented in Fig. 3.

**Algorithm 1** Enhanced EM based CCP Model

**Input:** CHs (σ)

**Output:** CH (σ), covariance matrix (Σ), and the nodes belonging to the individual cluster

𝑁 𝐶

𝖻 = 𝑙𝑛 𝐸(X|𝜎, Z, 𝜎) = ∑ 𝑙𝑛 {∑ 𝜎𝑐𝐴(𝑥𝑛|𝜎𝑐, Z𝑐)}

(4)

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

𝑛=1

𝑐=1

Estimate clustering parameters (σ)and Σ.

This process continues till convergence. The value of log likelihood is decreased monotonously that terminates EM process. As, EM updates key CH information including locationσc, and NR of each node connected to the cth cluster that eventually leads gradual reduction in SoS of the distances between each node and cluster that eventually gives optimal result.

1. *Proposed Enhanced EM Based CCM Clustering*

The presented research intends to develop energy efficient and QoS oriented data gathering protocol for densely distributed large scale WSN. Considering large volume of the sensor data, node-grouping is performed. As stated in previous section, the term “Group” refers the set of nodes who can communicate with each other. Here, a node can only communicate with the other nodes in the same “Group”. In fact, more number of clusters cause higher energy consumption, hence estimating optimal number of clusters to collect data from associated CNs can be vital. This research work targets on identifying the optimal number of clusters to perform efficient data gathering. In the proposed model, at first the mobile sink decides the CH arbitrarily at certain location. Employing the random location vector of the CHs, the communication distances Dncbetween each node and the connected CH has been estimated. In addition, the mixing coefficient, σ and Σ are also estimated, and thus initiating clustering, our proposed model selects group g using (5).

Calculate the parameters, Dnc and E.

***while*** |E − Enew| < γ ***do***

With the highest Sgvalue, perform “Group (g)selection”.

***for*** c ∈ Cg ***do***

***for*** n ∈ Ng***do***

Estimate RF value (φnc)for the nth node

## end for

Estimate the total number of nodes Nc belonging to the cluster,

Update the parameters σ, and Σ, using Nc.

## end for

Estimate log likelihood Enew.

## end while

Return CH (σ) , covariance matrix (Σ) , and the nodes belonging to the individual cluster.

Fig. 3. Enhanced EM (EEM) clustering for single mobile sink based WSN data gathering

1. ENHANCED EM (EEM) BASED CCP MODEL FOR

WSN DATA GATHERING

Once performing above mentioned EEM clustering, our model executes mobile sink movement to patrol each CH in the network so as to collect their data. Typically, the connected sensor nodes (CNs) transmit their data to the CH which later transmits data to the mobile sink (in case of mobile sink based data gathering); however it causes multi-

𝑆g

= 𝐶g

𝑁g

(5)

hop communication scenario that unsurprisingly can cause energy exhaustion and delay. The delay can be a type of waiting period between data generation, data transmission

where, Cg and Ng signify the total number of clusters in the group and the number of nodes in the group, respectively. In our model, the connected sensors or CNs from the group having highestSg, from that group and updates NR for all CNs. In our model, NR signifies the maximum extent to which a node n is connected to the clusterc. Now, using updated NR, the parameters σ and covariance matrix (Σ) are re-estimated. Eventually, the number of nodes belonging to the kth cluster are obtained as (6):

from CNs to CH and from CH to the mobile sink. As, the speed of mobile sink is slower than the electrical communication between CNs and hence there could be higher delay incurred because of single mobile sink node. To alleviate such issue, minimization of the patrolling distance can be significant, which can enable mobile sink to reach CH early to collect its data. To achieve this, a heuristic mode applying Traveling Salesman Problem (TSP) is applied that calculates minimum patrolling distance to reach CH. Reaching CH, mobile sink collects data from each connected CH through CH. Here, we have applied a variant of Directed Diffusion [32] technique called “One Phase Pull approach

[33] to perform data collection. In this scheme, the mobile sink sends a transmission request message (TRM) to the CHs in the network and getting TRM acknowledgement from cluster k, mobile sink re-transmits the data request to the CH, which is followed by relaying of the data from CH to the mobile sink. 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 time efficient data gathering the relative distance between CN and the CH or the mobile sink could not be addressed. Here, an additional scheduling measure has been incorporated 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.

To minimize the total energy needed for data transmission, all CNs transmit their sensed data as per the RF value of the cluster. In this work, RF value is calculated based on the network parametersμ, σ, andΣ, as discussed above (3). We have appended these parameters to the TRM transmitted by the mobile sink. Considering computational complexity of the proposed model, in our routing approach once deploying CNs, each node exchanges its location vector x with all CNs belonging to the same group. We perform information exchange of the location vector x only once and thus it avoids the unwanted signaling overheads and associated energy consumption.

As stated above, with the goal to reduce energy consumption, in this paper it is intended to estimate the optimal number of clusters to achieve energy efficient data gathering. The proposed cluster estimation model is given in the following sub-section.

*A. Reliabulity and Energy Efficint Cluster Formation*

Literatures [34] state that more clusters might reduction

other; else CM is updated with 0. It should be noted that to estimate the optimal number of clusters in the network, mobile sink knows the location of each nodes and respective communication range through which it can estimate the connectivity parameter using equation (7).

1. ENERGY EFFICIENT TRM FLOODING CONTROL

As stated in above discussion, reducing signaling overheads and packets can significantly reduce the energy consumption to enable optimal WSN data gathering. Once mobile sink reaches CH, it transmits a beacon message called TRM to call up for data transmission from CNs. Once CN gets the TRM, it transmits field or sensed data to the mobile ink through CH and broadcast a TRM to the neighboring nodes and TRM continues till all CNs in the cluster has received it. Practically, it can be feasible that a CN may get TRM twice or even more. CN transmits data first and then broadcasts TRM only after the first time of receiving the transmission request. No doubt, such uncontrolled transmission might cause energy exhaustion due to excessive redundant packet transmission. To avoid it, reducing or controlling redundant transmission overheads or TRM is must. Typically, TRM flooding might undergo higher connectivity requirements and hence could introduce huge signaling overhead. In addition, increase in TRM results into the increases in the number of clusters and hence requires capping TRM counts and minimal clusters to ensure reliable as well as energy efficient data gathering. To achieve this, we have defined the sum of required energy for TRM as the objective function 𝑍(𝐶) to be minimized that can be obtained as the sum of energy consumption occurred due to the mobile sink patrolling (each cycle). Mathematically,

Z(C) = LReqSReq(C) + LDatSDat(C), (8)

where 𝑆𝑅e𝑞(𝐾) and 𝑆𝐷𝑎𝑡(𝐶) presents the SoS of transmission distance of data requests and data messages, respectively. The parameters LReq and LDat present the size of the data and the TRMs, respectively. Mathematically, SDat(C) is obtained as (9).

in the energy consumption; however, majority of the

approaches don’t address energy consumption issues caused

𝑁 𝑐

𝐻𝑛𝑐

(9)

due to signaling process (particularly TRM signaling caused energy consumption). To deal with this issue, we propose a novel cluster count estimation model that intends to keep minimum number of clusters while maintaining higher throughput and timely data delivery. To enable reliable and QoS oriented transmission, the inter-relationship between network connectivity and associated energy consumption is examined. To assess the correlation between CNs and mobile sink connectivity, a connectivity assessment model is derived where connectivity is defined as a fraction of nodes that could communicate with each other. Mathematically (7).

𝑆𝐷𝑎𝑡 = ∑ ∑ ∑ 𝜑𝑛𝑐. 𝑙2

𝑛=1 𝑐=1 =1

where Hnc refers the total number of hops from nth node to the cth CH and lh presents the communication distance of each hop. In case nth node is unable to communicate with cth CH, Hnc is set as 0, and thus the energy required is 0. Here, SDat(C) is the decreasing function of C while TRM energy SDat(C) is often increasing function of C. It signifies a trade-off correlation in between the first and second component of (8). Now, to maintain K more than the total number of groups G, the optimal number of clusters Coptimal

CM =

G

g=1

∑

Ng(Ng

− 1)

(7)

is obtained by (10)

( ) (10)

N(N − 1)

Here, CM refers binary value (0 or 1), in which CM=1 refers a condition when all nodes can communicate with each

𝐶o𝑝𝑡i𝑚𝑎𝑙 = 𝑚𝑎𝑥 (𝐺, 𝑎𝑟𝑔 𝑚i𝑛(𝑍 𝐶 ))

𝑐

Now, to calculate the required transmission power for TDMs, the set of nodes in a group Ng and number of cluster

heads Cg are considered. Here, TDM is transmitted from the individual CH and each CN retransmits it once and thus the total energy required for TDM transmission is (11).

(distance between sensor node (CN) and mobile sink or CN and CH) can be effective to avoid data drop probability and retransmission. Thus, the proposed transmission model can

𝐺

𝑆𝑅e𝑞 = ∑ 𝐶g𝑁g𝑀2,

g=1

(11)

be effective solution for load sensitive transmission

scheduling for WSN data gathering.

VII. RESULTS AND DISCUSSION

where M refers the maximum radio range of CNs. In case of non-imbalance of the CHs location, CNs belonging to each cluster is equivalent. i.e.,

Considering effectiveness of expectation maximization (EM) based centralized clustering model (CCM) for WSN data transmission, in this paper an enhanced routing model

𝐶g

𝑁g

= 𝐶

𝑁

(12)

is developed. As proposed, this research intended to incorporate multiple constraints optimization measure to achieve a better load sensitive, delay sensitive and QoS oriented data gathering model, where the proposed enhanced

With multiple CN, CM is approximated as (13) EM model (EEM) based CCM has outperformed other

g=1 g g

𝐶𝑀 =

∑𝐺 𝑁 (𝑁

− 1)

𝐺 2

o g=1 g

∑ 𝑁

(13)

existing centralized clustering approaches such as K-Conid [36]. To further optimize, an additional constraint has been

𝑁(𝑁 − 1)

𝑁2

introduced on the basis of the distance between CNs and the CH as well as CNs and mobile sink, so as to enable time-

Thus, applying (11-13), SReq is obtained as (14).

𝑆𝑅e𝑞 = CN ∗ M2C (14)

Equation (14) presents that the energy required for TRM transmissions is directly proportional to the connectivity and as the number of clusters affects the connectivity it eventually affects the energy consumption. In addition it (14) refers that the energy consumption increases as per the total number of clusters C and hence lower cluster counts can be anticipated to enable minimal energy exhaustion for efficient data gathering. Thus, measuring the energy required for data gathering at the mobile sink (14), and TRM (9), the optimal clusters count can be obtained (10).

1. INTER SENSOR NODE-CH AND MOBILE SINK DISTANCE BASED LOAD SENSITIVE DATA GATHERING 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. Considering load sensitive data gathering, in case a CN possessing large data content can schedule its transmission model based on traffic condition and inter-node distance factor. Transmitting data with minimal distance

efficient WSN data gathering. Applying inter-node distance (𝑑i𝑠𝑡(𝐶𝑁𝑠𝐶𝐻𝑀o𝑏i𝑙e 𝑆i𝑛𝑘)o𝑟 CNs Mobile Sink) based dual transmission strategy enables optimal data transmission or gathering. Unlike existing systems where traditional clustering models are used, we developed an each-node’s degree of dependence (DoD) based “Node Responsibility (NR) factor that effectively enhances generic EM based clustering to assists data sensitive reliable transmission. In addition, with the goal to reduce the number of clusters in the network, a reliability and energy efficient cluster formation scheme is developed.The dual transmission stratgy employs two ways for data gathering, in which at first CN transmits data to the CH to which it belongs, which is then transmitted from CH to the mobile sink. In the second 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. If a CN finds mobile sink nearer to the CH, it transmits its data directly to the mobile sink that significantly reduces the computational overheads, delay, energy exhaustion and relaying cost etc. The overall ruting model for data gatehring is developed using Network Simulator (NS) ver. 2, usually called NS2. To enable better result presentation, MATLAB software has been used to plot graphs. The experiemental simulation environment is presnted in Table 1.

TABLE I. 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) |

Once performing cluster formation and introducing mobile sink based data gathering mechanism; the energy consumption for the data transmission (EDat)from CNs to the mobile sink was estimated. Here, EDat signifies the amount of energy required to perform data transmission from sensor nodes to the mobile sink. Here, in case the position of CH is far away and it can’t make connection with CH, EDatvalue is updated with 0 that reflects failure of clustering and doesn’t mean the energy conservation. The research simulation with the generic EM based clustering doesn’t consider any connectivity issue undergoes degraded performance. On the other hand, K-Conid [29] based clustering that applies generic EM based clustering for data transmission is considered as a reference model. However, as our proposed model, we have at first applied EEM clustering to perform data gathering. We have examined our proposed model in two ways. In first approach, with the EEM based clustering followed by optimal cluster count estimation (OCCE) and transmission in “CNsCHMobile Sink”. 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.



Fig. 4. Packet delivery ratio

Fig. 4 presents the performance of the proposed data gathering algorithms in terms of the packet delivery ratio (PDR). Here, it can be observed that EEM model which is the enhanced version of the generic EM clustering performs satisfactory. Interestingly, the data gathering model with (CNs Mobile Sink) outperforms “CNsCHMobile Sink” transmission. 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. This can be due to additional losses incurred due to receiving and transmission at the CH node. The relative distance between CN and the 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 (Fig. 4) 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 Fig. 7, 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 CHs is must to ensure higher throughput and reliable data gathering. 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 Fig. 5. 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 CN to the CH and then from CH to mobile sink. Transmitting data from CN to the mobile sink can reduce the probability of packet drop and unwanted traversal that eventually can lead reduced end-to- end time in data gathering. Fig. 5 exhibits that the proposed transmission model can perform fast data gathering, which is of paramount significance to ensure delay sensitive communication over WSNs.



Fig. 5. Transmission or data gathering delay

Considering energy efficiency, Fig. 6 depicts that the proposed data gathering model 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.



Fig. 6. Energy consumption by different techniques

For any routing protocol, its computational efficiency plays a vital role to enable it for real world applications. Here, we have examined the performance of the proposed routing protocol in terms of efficiency. Considering aforementioned discussions it is visible that those clustering approaches avoiding connectivity issue undergo 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):

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 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

𝐸ffi𝑐ie𝑛𝑐𝑦 = (𝑁𝑢𝑚𝑏e𝑟 of 𝑐o𝑛𝑛e𝑐𝑡e𝑑 𝑛o𝑑e)

𝑆𝐷𝑎𝑡

(15)

a cluster due to unavailability of mobile sink.

The performance assessment of the proposed data gathering scheme with existing distributed clustering K- Conid [36] and our proposed CCP based clustering and data gathering, it can be found (Fig. 7) that the proposed data gathering approach exhibits higher efficiency than the other approaches.



Fig. 7. Network communication efficiency

VIII. CONCLUSION

The exponential rise in the communication system demands have motivated academia-industries to develop certain low-cost, energy efficient 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

REFERENCES

1. E. B. Hamida, and G. Chelius, “Strategies for data dissemination to mobile sinks in wireless sensor networks,” IEEE Wireless Communications, vol. 15, no. 6, pp. 31–37, 2008.
2. F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, “A two-tier data dissemination model for large-scale wireless sensor networks,” in Proceedings of The 8th Annual International Conference on Mobile Computing and Networking (Mobicom ’02), pp. 148–159, 2002.
3. M. Taifeh Aliverdi, and M. R. Pakravan, “A Network information gathering (NIG) protocol and energy-aware load balancing algorithm for wireless sensor network,” Telecommunications and Malaysia International Conference on Communications, ICT-MICC, IEEE International Conference on, Penang, pp. 407-412, 2007.
4. K. Almi’ani, A. Viglas, L. Libman, “Energy-efficient data gathering with tour length-constrained mobile elements in wireless sensor networks,” 35th Annual IEEE Conference on Local Computer Networks LCN 2010, Denver, Colorado, 2010.
5. L. Cheng, Y. Chen, C. Chen, J. Ma, “Query-based data collection in wireless sensor networks with mobile sinks,” International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly, pp. 1157-1162, 2009.
6. Z. Sha, J. Lu, X. Li, M. Wu, “An anti-detection moving strategy for mobile sink,” IEEE Proceeding of Global Telecommunications Conference (GLOBECOM 2010), 2010.
7. L.Cheng, Y. Chen, C. Chen, and J. Ma, “Query-based data collection in wireless sensor networks with mobile sinks,” International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly, pp. 1157-1162, 2009.
8. A. Kinalis, S. Nikoletseas, D. Patroumpa, J. Rolim, “Biased sink mobility with adaptive stop times for low latency data collection in sensor networks”, STREP ICT-257466, Information Fusion Journal, pp. 1566-2535, 2012.
9. T. Park, D. Kim, S. Jang, S.Yoo, Y. Lee, “Energy efficient and seamless data collection with mobile sinks in massive sensor networks”, Parallel and Distributed Processing, 2009.
10. Z. Sun, Z. Zheng, S. Xu, and S. Chen, “Energy-aware data gathering protocol for wireless sensor networks,” Computer Science and Information Technology (ICCSIT), 3rd IEEE International Conference on, Chengdu, pp. 103-107, 2010.
11. A. Boukerche, and R. W. Nelem Pazzi, “Lightweight mobile data gathering strategy for wireless sensor networks,” Mobile Wireless Communications Networks, 2007 9th IFIP International Conference on, Cork, pp. 151-155, 2007.
12. D. Lee, S. Park, E. Lee, Y. Choi, and S.-H.Kim, “Continuous data dissemination protocol supporting mobile sinks with a sink location manager,” in Proceedings of the Asia-Pacific Conference on Communications (APCC ’07), pp. 299–302, October 2007.
13. S. Nesamony, M. K. Vairamuthu, and M. E. Orlowska, “On the traversals of multiple mobile sinks in sensor networks,” in Proceedings of the IEEE International Conference on Telecommunications and Malaysia International Conference on Communications (ICTMICC ’07), pp. 432–437, May 2007.
14. L. Song and D. Hatzinakos, “Dense wireless sensor networks with mobile sinks,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’05), vol. 3, pp. 677–680, March 2005.
15. Z. M. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli, “Exploiting sink mobility for maximizing sensor networks lifetime,” in Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS ’05), January 2005.
16. P. Traynor, J. S. Shin, B. Madan, S. Phoha, and T. La Porta, “Efficient group mobility for heterogeneous sensor networks,” in Proceedings of the IEEE 64th Vehicular Technology Conference (VTC’ 06), pp. 1–5, September 2006.
17. Q. Qiu and A. E. Kamal, “Coverage and connectivity control of wireless sensor networks under mobility,” in Workshop on High Performance Switching and Routing (HPSR ’05), pp. 177– 181, May 2005.
18. S. Guo, C. Wang, Y. Yang, “Joint Mobile Data Gathering and Energy Provisioning 184 in Wireless Rechargeable Sensor Networks,” IEEE Transactions on, Mobile Computing, vol. 13, no. 12, pp.2836-2852, Dec. 2014.
19. M. Ma, Y. Yang, M. Zhao, “Tour Planning for Mobile Data- Gathering Mechanisms in Wireless Sensor Networks” IEEE Transactions On Vehicular Technology, Vol. 62, No. 4, May 2013.
20. D. Jea, A. Somasundara, and M. Srivastava, “Multiple controlled mobile elements (data mules) for data collection in sensor networks,” in Distributed Computing in Sensor Systems, vol. 3560, pp. 244–257, 2005.
21. D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, “Toward Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks,” In. IEEE Trans. Emerging Topics in Computing, pp. 388- 397, 2014.
22. P. Madhumathy, and D. Sivakumar, “Enabling energy efficient sensory data collection using multiple mobile sink,” in China Communications, vol. 11, no. 10, pp. 29-37, Oct. 2014
23. C. Zhu, S. Wu, G. Han, L. Shu, H. Wu, “A Tree-Cluster-Based Data- Gathering Algorithm for Industrial WSNs With a Mobile Sink,” in Access, IEEE , vol. 3, pp.381-396, 2015.
24. W. Wang, V. Srinivasan, and K. C. Chua, “Using Mobile Relays to Prolong the Lifetime of Wireless Sensor Networks,” In: Porta TL, Lindemann C, eds. Proc. of the 11th Annual Int’l Conf. on Mobile Computing and Networking. New York: ACM Press, pp. 270-283, 2005.
25. C. Gherbi, Z. Aliouat, and M. Benmoh ammed, “Distributed energy efficient adaptive clustering protocol with data gathering for large scale wireless sensor networks,” Programming and Systems (ISPS), 12th International Symposium on, Algiers, pp. 1-7, 2015.
26. W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energyefficient communication protocol for wireless microsensor networks,” in Annual Hawaii International Conference on sytem Sciences, vol. 2, Jan. 2000.
27. R. C. Shah, S. Roy, S. Jain, and W. Brunette, “Data MULEs: modeling and analysis of a three-tier architecture for sparse sensor networks,” Ad Hoc Networks, vol. 1, no. 2-3, pp. 215 – 233, 2003.
28. M. Youssef, A. Youssef, and M. Younis, “Overlapping multihop clustering for wireless sensor networks,” IEEE Transactions on Parallel and Distributed Systems, vol. 20, no. 12, pp. 1844–1856, Dec. 2009.
29. T. Khac and C. Hyunseung, “Connectivity-based clustering scheme for mobile ad hoc networks,” in IEEE International Conference on Research, Innovation and Vision for the Future, Jul. 2008.
30. S. Lindsey and C. Raghavendra, “PEGASIS: Power-efficient gathering in sensor information systems,” in IEEE Aerospace Conference, 2002.
31. H. Nakayama, N. Ansari, A. Jamalipour, and N. Kato, “Fault-resilient sensing in wireless sensor networks,” Computer Communications, vol. 30, no. 11-12, pp. 2375–2384, Sep. 2007.
32. C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed Diffusion: a scalable and robust communication paradigm for sensor networks,” in MobiCom’00 Proceedings of the 6th annual international conference on Mobile computing and networking, 2000.
33. M. Chen, T. Kwon, and Y. Choi, “Energy-efficient differentiated directed diffusion (eddd) in wireless sensor networks,” Computer Communications, vol. 29, no. 2, pp. 231–245, 2006.
34. D. Takaishi, H. Nishiyama, N. Kato, and R. Miura, “Towards Energy Efficient Big Data Gathering in Densely Distributed Sensor Networks,” IEEE Transactions on Emerging Topics in Computing, 2168-6750, 2013.