International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 1st Jun 2019. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-06

Title: LINK-AWARE DATA-BASED CLUSTER-TREE FORMATION SCHEME WITH MULTI-SINK TO IMPROVE NETWORK PERFORMANCE IN LARGE-SCALE WSN

Volume 08, Issue 06, Pages: 209–219.

Paper Authors

PALLAVI A, SHILPA T, DR R NAGARAJA

Bangalore Institute of Technology, Bangalore

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

LINK-AWARE DATA-BASED CLUSTER-TREE FORMATION SCHEME WITH MULTI-SINK TO IMPROVE NETWORK PERFORMANCE IN LARGE-SCALE WSN

[1]PALLAVI A, [2]SHILPA T, [3]DR R NAGARAJA
1 PG Student (M.Tech), Department of ISE, Bangalore Institute of Technology, Bangalore
2 Assistant Professor, Department of ISE, Bangalore Institute of Technology, Bangalore
3 PG and Research Coordinator, Department of ISE, Bangalore Institute of Technology, Bangalore

Abstract—In the Large-scale Wireless Sensor Network there is a difficulty in transferring huge amount of data to destination i.e. Base Station. While transmitting data there may be large energy consumption, latency, which directly affects the network performance. Link quality may reduce due to node failures, traffic, and reduced network lifetime. To overcome these challenges a hierarchical top down cluster tree formation scheme is used called Data-based Cluster Tree Formation (DbCTF).This DbCTF scheme forms a Data Cluster Tree (DCT) using the Data Collecting Node (DCN), a type of information gathering node attached to each Cluster Head (CH) from one hop distance, which ultimately forms a cluster tree to the Base Station. In this paper the DbCTF scheme with a DCN node for multiple Base Stations is proposed. Simulation is done using NS-2.34 simulator, which demonstrates that DbCTF scheme provides better Quality of Service (QoS) in terms of throughput, end-to-end delay which is called latency, packet delivery ratio (PDR) including overhead. By comparing the simulation results for single and multiple Base Stations, it is relieved that for multiple base stations yields better results in terms of average network lifetime.

Keywords— Data-based cluster tree, data collecting node, data collection tree, wireless sensor networks.

I. INTRODUCTION

Large-Scale Wireless Sensor Network is a type of wireless network; this network usually contains numerous various types of sensor nodes. In which nodes are deployed randomly in the monitored region provided with divergent energy. Every sensor node in the wireless network consists of separate sensing unit, processing unit, storage and communication unit. It is not possible to predetermine the position of the sensor node. This allows the nodes to deploy randomly in the network. Sensor nodes are characterised with low power and small range, hence, these nodes must work together in multi-hop wireless communication systems that allow the transmission of sensed data and collect information to the nearby sink. The first development in WSN was motivated by its various applications namely military uses to do surveillance in battlefield. The advancement of new technologies decrements the size, cost and power of sensor nodes to make Wireless Sensor Network one of the major topics of

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

wireless communication. In any Wireless Sensor Network there are four basic components namely, an interconnecting wireless network; a collection of scattered sensor nodes; a collecting data sink; a collection of computation equipment at the sink or for away the base station for analysis and interpretation of received data from the nodes. Figure 1 shows the WSN components.

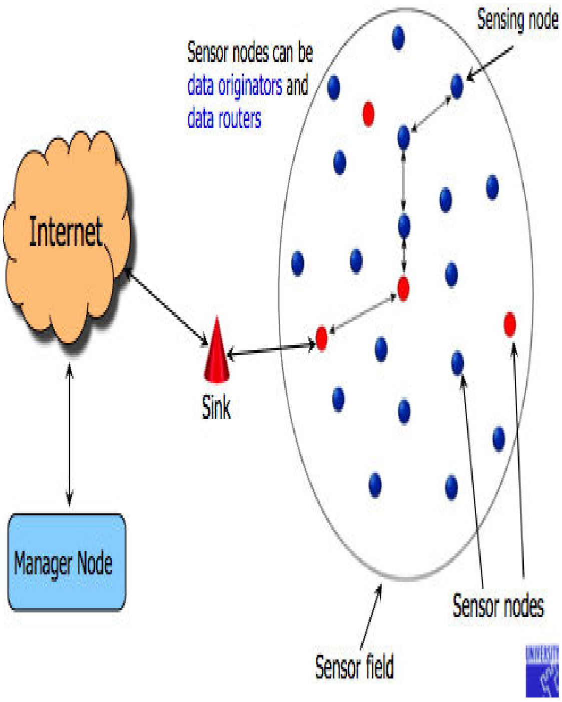

Figure 1: WSN Components

Sensor nodes are useful for accumulation of specific type of data and transfer it to the base station as mentioned earlier. Each sensor node contains four components as shown in Figure 2, in which the nodes are furnished with a sensing unit, a radio transceiver or other wireless communication gadgets, a small control tool which incorporates a microprocessor, and an energy source, and also additional memory component is there in some sensor nodes.

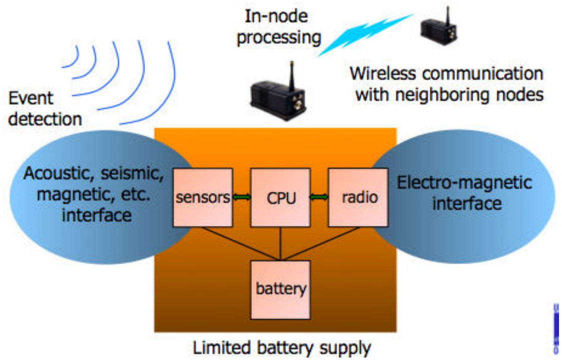

Figure 2: Additional components in WSN

Functionality of the sensor nodes actually relay on the network topology structure that in turn rest on utilization. Hence its functionality lies behind the node's ability to either being the source of the data and then transmits it. In order to reach the base station it just became a transceiver that receives data from other sources and then transmits it to other nodes. Table 1 represents the specification of each node.

Table 1: List of Sensor Nodes

Sensor Node Name	Microcontroller	Tranceiver	Program+Data Memory	External Memory
BEAN	MSP430F169	CC1000 (300-1000 MHz) with 78.6 kbit/s		4 Mbit
BTnode	Atmel ATmega 128L (8 MHz @ 8 MIPS)	Chipcon CC1000 (433-915 MHz) and Bluetooth (2.4 GHz)	64+180 K RAM	128K FLASH ROM, 4K EEPROM
COTS	ATMEL Microcontroller 916 MHz			
Dot	ATMEGA163		1K RAM	8-16K Flash
EPIC Mote	Texas Instruments MSP430 microcontroller	250 kbit/s 2.4 GHz IEEE 802.15.4 Chipcon Wireless Transceiver	10k RAM	48k Flash
Eyes	MSP430F149	TR1001		8 Mbit

The architecture of the network is directly related to the application deploying wireless sensor network. For example some of the nodes in the network are directly connect to the base station without passing it through other nodes (1-hop layer).Other layers may pass through other nodes to transfer data to sink. Figure 3 shows these different layers.

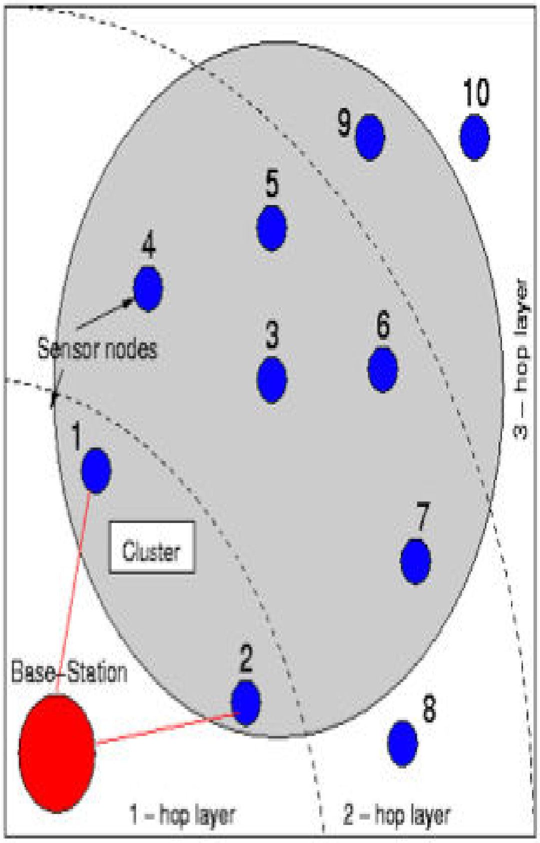

Figure 3: WSN Architecture with Layers

The protocols of IEEE are used broadly and they are implemented in Wireless Sensor Network technology, there are various protocols in WSN; The most widely used are 1.IEEE 802.11(WLAN); 2.IEEE802.15.4 (ZigBee); 3.IEEE802.15.1 (Bluetooth).

Network topography authority plays a significant part in minimizing several challenges namely nodes damages, long-distance communication within a network, delay, traffic, failure in communication etc. The enhancement of QoS in the network, performance, coverage and networks lifetime is done by selecting a right topology. The efficient topology ensures that it decreases the chance of a packet being stray between sensor nodes and also confirms the neighbours are at a minimal distance and. Energy Consumption is the main variable plays a significant part in act of WSN's. Energy Consumption is straight away linked to transmission distance between the sensor nodes in the wireless network.

II. RELATED WORK

Generally the network topology in WSN to dissolves the universal effectiveness of the network. Various algorithms are proposed to overcome challenges in WSN, some of them are 1. A Clustering-tree topology control algorithm based on the energy forecast (CTEF); 2. The Tree-Cluster based data-gathering algorithm (TCBDGA) for WSN; 3. Velocity Energy-efficient and link-aware cluster tree (VECT) algorithm.

A. *CTEF Algorithm*

Clustering-tree topology control algorithm based on energy forecast (CTEF) Algorithm for Heterogeneous Wireless Sensor Network (HWSN), since HWSN is not energy-efficient and equipped with minimum network lifetime. CTEF Algorithm resolves the challenge of more energy consumption. The CTEF is mainly used for energy saving and ensuring load balancing in the network. By using concurrently both the Central limit theorem and normal distribution, the average energy of the network per round is accurately predicted by using CTEF strategy .Selection of Cluster Head (CH) is done using Cost function and their distance. This algorithm searches for the distributed nodes in the cluster to forwards information and also to decreases the CH burden. The outcome obtained by simulation shows CTEF method is effective only for HWSN and its actions in case of lifetime and throughput is better than LEACH and EDFCM protocols. However for large-scale WSN CTEF fails to achieve maximum average network lifetime and energy efficiency.

B. *TCBDGA Algorithm*

TCBDGA for WSN is proposed to overcome the Hotspot problem, which affects the network connectivity and reduces the network lifetime. In this algorithm root nodes in constructed trees are defined as Rendezvous points (RPs) and in addition few peculiar nodes called Sub rendezvous points (SRPs) are elected by considering traffic load and hops to root nodes. Data gathering in WSN can be categorized as Direct Communicating Collection and Multi-hop Communicating Collection (MCC). Depends on the contents of data need to be collected MCC can be divided into two sub categories. They are

1. Local Data Collection and
2. Full data Collection

In the proposed TCBDGA, the base station initiates the collection of data periodically from the base station and it stops at every RP, captures information directly from nodes in its one-hop range, at the end return back to base station for one round. The entire TCBDGA method includes mainly three steps namely Tree Formation; RPs and SRPs selection and Data Collection. Adjusting method is introduced for RPs and SRPs reselection. After simulating, results demonstrate that TCBDGA scheme can prolong the lifetime of the network remarkably. But it is suitable only industrial environment containing large number of various sensory data.

C. *VECT Algorithm*

The lifetime of sensor nodes is extended by using data aggregation protocols. In event-based applications the structured approaches can match the data gathering applications, they also incur high maintenance overhead. The VELCT method minimises end-to-end delay, maximises energy exploitation and also reduces traffic in clustered WSN by effective usage of DCT. In order to avoid cluster formation frequently a simple tree structure was constructed, which also helps in decreasing energy consumption of the cluster head and this is considered as the strength of VECT method. The identified network topologies are Flat topology; Chain-based topology; Cluster-based topology; and Cluster Tree-based topology (CTT). The VELCT strategy is proposed to construct network management architecture for WSN. Similarly, the sink or Data Collecting Node finds the one-hop neighbour DCN node or CH with less network traffic.

The VELCT mechanism mainly contains two phases: Set-up phase and Steady-state phase. The set-up phase initiates the cluster formation and DCT construction process for identifying route between cluster member and sink. Then in steady-state phase data transfers from cluster members to the sink. VELCT provides

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

suitable links and PDR with minimum network traffic than existing protocols.

III. PROBLEM STATEMENT

In sensor nodes communication is the main factor of energy decadence. The dissolution relay on the distance between communicating elements such as source and destination. Every sensor node in the network are able to communicate directly with the base station or it just forward the packets containing data to neighbour nodes which are at one-hop distance and finally reaches the base station. Limitations in the existing methods are latency, failure of nodes, redundancy in data and more utilization of energy, because its uses flooding and direct communication method for nodes communication. It is the main drawback of existing topologies which is not recommended for large-scale wireless sensor network. Chain topology also has some demerits namely, communication delay, failure of nodes and exploitation of energy. Disseminating the data from one cluster head to another cluster head or to base station either by direct-hop or multi-hop requires stable links that are reliable, it leads to huge amount of energy consumption. The responsibility of sensor nodes in multi-hop communication is receiving and forwarding data from the entire network to the nearest base station, this tends to consume more energy, which results in non-uniform energy consumption between nodes, which are called as “hotspot problem”. In sensor network this results to link failure and transmission delay. The above mentioned topologies became probably infeasible and cannot be adapted to large-scale wireless sensor network. Cluster Tree Topology is a hybrid scheme compared with above mentioned topology structures. Next section describes how the DbCTF scheme is implemented with single base stations and with multiple base stations.

IV. THE DbCTF DESIGN

This section describes the Data-based Cluster tree formation scheme named DbCTF. The existing wireless sensor network topology structures includes flat, tree, cluster, chain and hybrid types. Various topology structures are followed to meet the maximum data collection and send it to the sink based on the network nature. Tree topology structure can reduce more energy consumption compared to existing flat and chain topology. Hybrid topology gives better performance compare to available single topologies. Initially for cluster tree creation a state-of-the-art approach is presented. It is called Baseline Algorithm.

A. *Baseline Algorithm*

Cluster-tree network is formed by the Baseline algorithm by considering the only criteria the connectivity among nodes. The process of forming network is carried out in a top-down approach, in which Base station takes the responsibility to initialize the network formation process. Nodes send request to the CH after receiving the invitation frame to associate. Considering particular criteria like maximum number of nodes per cluster. The cluster head sends acknowledgment messages to all accepted association requests and negative acknowledgements to others. This recursive process will be done until all nodes are associated with the network.

The main intension is to maximise the coverage of cluster tree, with maintenances of link quality between nodes and CHs.

B. *DbCTF Strategy for single sink*

The Data base Cluster tree formation scheme works using following steps

1. Random Node Distribution
2. Cluster Formation and Cluster Head Selection
3. Data Collection Tree (DCT) construction
4. Data aggregation and Transmission
5. Performance Evaluation.

Objective of the proposed system is as follows:

- To find the energy efficient data collection scheme to increase the network lifetime.
- To implement data collection nodes in each cluster, to help cluster head in forwarding data to sink.
- To search the optimal and energy efficient multipath to reach destination.
- Make opportunistic routing to achieve high energy efficiency of the network.

Advantages of the DbCTF scheme are as follows:

- Less Delay
- Computation overhead is less
- More network lifetime
- Less energy consumption

The overall System architecture is as shown in Figure 4.

System Architecture:

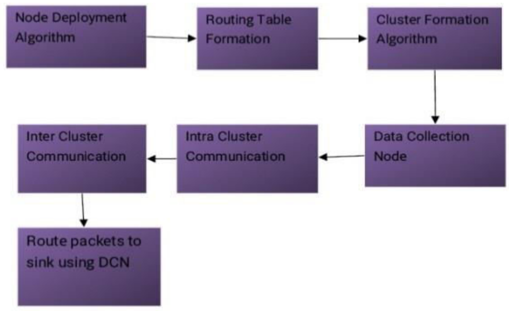

Figure 4: System Architecture

The brief explanation of each steps are as follows:

The first step is the distribution of nodes randomly in sensor network

Algorithm 1: Node distribution algorithm

The pseudo code for initialization of nodes randomly is as follows

1. for i=0 to num_of_nodes
2. Initialize mobile nodes
3. Enable random motion
4. Nodes Energy
5. Define initial node position
6. Initialize agent
7. Attach agent to node
8. End for

The second step is the most important one that is Cluster Formation and Selection of Cluster Head (CH). To construct the topology cluster-tree design is used. In the wireless sensor network nodes helps to form clusters each cluster includes a CH and Cluster Members (CMs) which are located within the transmission range of corresponding CH. Clustering is done by

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

taking One-hop distance. The Clustering in WSN in as shown in below Figure 5.

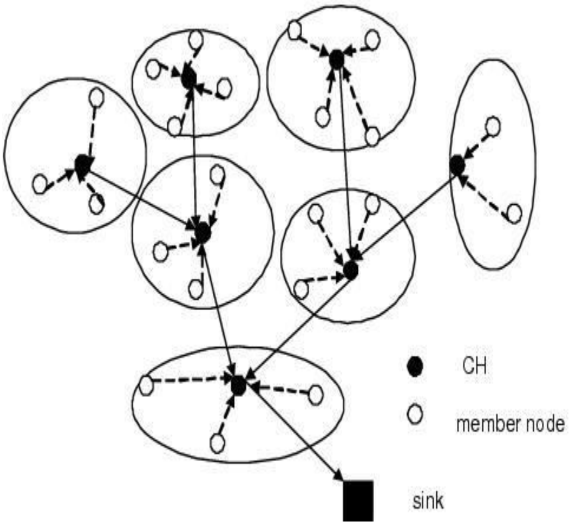

Figure 5: Clustering in WSN

Cluster Formation algorithm partition the whole network into different zones, in which each cluster consists of group of nodes in its zone. The algorithm is responsible for node deployment. Each zone is bounced with limits contains some xmin and xmax. The y area is covered within the limits ymin and ymax.

Algorithm 2: Cluster Formation algorithm
Below steps represents the Cluster Formation algorithm

1. Input: Number of Clusters, Vertical and Horizontal end points for each cluster.
2. Repeat Step3, 4 and 5 until all clusters are formed for each Cluster Formation.
3. For the cluster select x end points and y end points appropriately
4. Cluster id should be generated for the cluster.
5. Deploy nodes randomly and place the nodes within the cluster.

Figure 6 describes how cluster member nodes send a request message to join with the selected CH. This request message contains node's ID and CH ID.

For Cluster Head selection the Residual Energy is calculated by Residual energy is calculated by using following equation:

Residual_Eg=Initial_Eg– Consumed_Eg

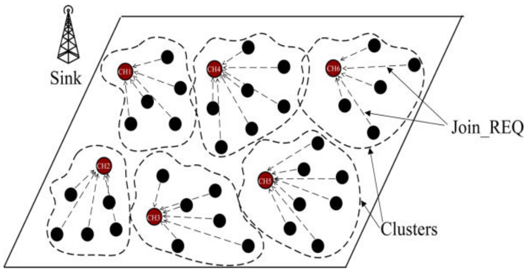

Figure 6: Cluster formation: node send request message to join.

After Cluster formation Cluster Head (CH) will be selected based on the High Residual energy, initially nodes are assigned with some energy, where residual energy is given by initial energy minus consumed energy. The sink node is responsible to select cluster head. The average residual energy of network is calculated by sink node and streaks only the nodes those are having energy higher than the average energy. Figure 7 shows the cluster head selection process in which sink node broadcasts the message containing CH node ID.

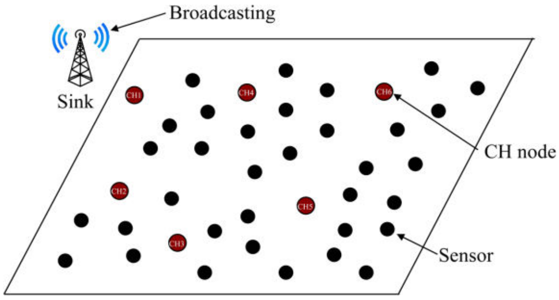

Figure 7: Cluster Head selection process

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

The next step is the construction of Data Collection Tree (DCT) using Data Collection Node (DCN). A DCN node can be defined as the following structure:

```
Struct Node
{
Pos_X; // X direction coordinate
Pos_Y; // Y direction coordinate
Select; // if the node adds a route
//data transmission energy consumption
// from node k to its neighbor node
DistanceToAdjacent[k];
//data transmission energy consumption
// from node to sink node
DistanceToSink[l];
Energy; // Residual energy
};
```

Data Collection Tree (DCT), an hierarchical tree structure, uses Data Collecting Node (DCN) for collecting the data from the cluster heads and deliver it to base station, which encloses entire wireless sensor network. Based on RSS (Received Signal Strength) and communication range base station selects the DCN, which minimizes surplus energy usage and traffic of the entire network. Whenever new cluster heads are elected, new DCN is selected every. New DCN selection is carried by sink, based on the new cluster head .mobility The DCN collects the data from cluster head, aggregates the data and then forwards it to the next DCN.

DCT Construction: Initially the sink starts to find a first DCN from one-hop distance neighbour nodes to add that particular node in DCT. The parameter includes Hop Count, HC=1 (i.e., HC is the Hop distance) is used to select the Current node identity that is one-hop distance Neighbour Node (NN) from sink or DCN. Sink or DCN verify parameters. After this, the sink selects a optimum node value from Current node identity and assigns the node integrity into TIN (true identity number). Now, the DCN is selected from the TIN. After DCN validation, node checks the network traffic. If the maximum frame duration is greater than traffic the elected DCN can used to form a DCT else, sink jumps to elect a new DCN and then starts to build a new DCT. Once the first DCN is selected, then the next DCN identification starts from the first DCN instead of sink. Consequently, the DCN selects another one-hop distance neighbour node to act as a current node identity. To generate the DCT to connect all of the cluster heads in the WSNs DCN selection is expanded. The data collected by the cluster members are sent to the cluster head. Then cluster head initialize to collect and aggregates data from its cluster members. At the same time, the DCT communication is initiated to transfer the data from the cluster head to DCN and then the sink. Finally the DCN collects and aggregates the data from the corresponding cluster head or DCN. Figure 8 represents the DCT formation in single cluster.

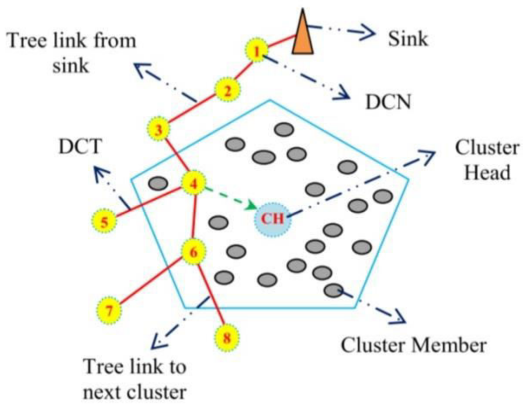

Figure 8: DCT construction in single cluster

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

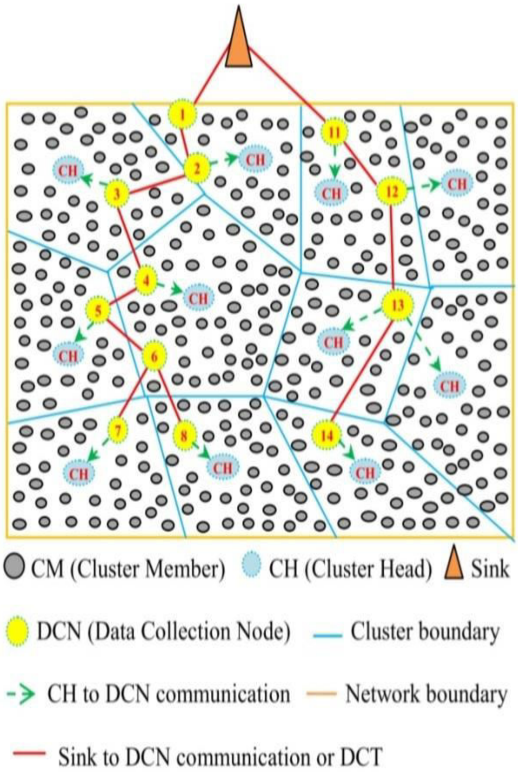

Figure 9: Tree formation among DCN nodes (Data aggregation and transmission)

The next step after the construction of DCT to the sink is the Data Aggregation and Transmission. Data transmission starts when sensor node sends data packets to their respective CH. The CHs then transmits the data to the corresponding DCN node, and then the DCN node aggregates data received from CHs and forwards them to the next higher level nodes. This is done until aggregated data reaches base station. Figure 9 mentions how sensor nodes forward sensed data to CH, and DCN aggregates and send them to sink.

The last step is the Performance Analysis

Performance evaluation of DbCTF scheme is carried out using NS- 2.34 and compared with LEACH, EDFCM, EDCS protocols. To do performance analysis the DbCTF strategy is implemented for multiple Base stations. A single sink scenario is not efficient when there is a lot of information to take care, then network density increases, furthermore a multi-sink environment is used. The main benefits of the multi-sink proposal are information redundancy, regulated energy distribution and load balance during tree formation. Figure 10 presents the result of the DbCTF scheme proposed for multiple sinks, where each sink forms tree.

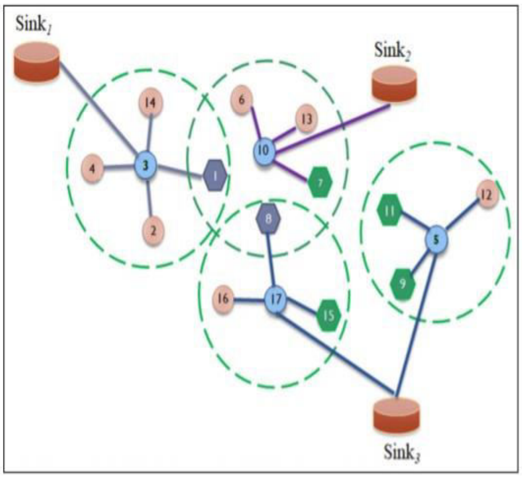

Figure 10: Tree formation in multi-sink.

Furthermore while simulating in NS-2.34 environment for single sink and multiple sink are looks like as shown in figure 11 and figure 12.

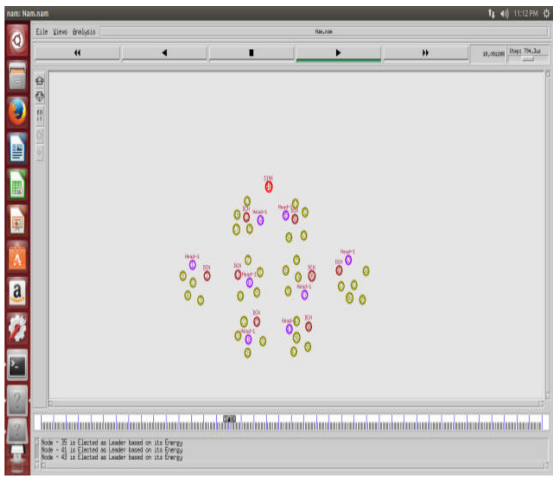

Figure 11: DbCTF strategy for single sink

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

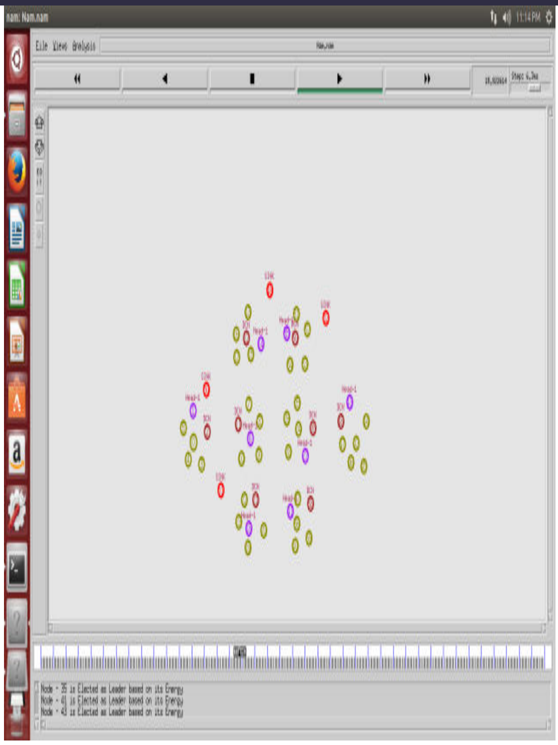

Figure 12: DbCTF strategy for multi-sink

V. RESULTS AND DISCUSSION

In this section performance of DbCTF scheme under various parameter settings via simulations is presented. The NS-2.34 is used to carry out the performance study of DbCTF with respect to existing topology. A WSN system consisting 50 nodes is used in the simulation. Graphs obtained after simulation by taking parameters like throughput, end-to-end delay, packet delivery function and overhead are shown in following figures. Based on the DbCTF, the network performance was simulated in terms of the total energy, delay, overhead, throughput and packet delivery ratio (PDR).Figures 13, 14 and 15 illustrate the relationship between the number of deployed nodes and the performance of the network.(total energy consumption ,delay, overhead).In the simulated graphs green line is for representation with single sink performance and red line represents the multi-sink. Figure 13 represents that total energy consumption for multiple sink is less compared to energy consumed for single sink.

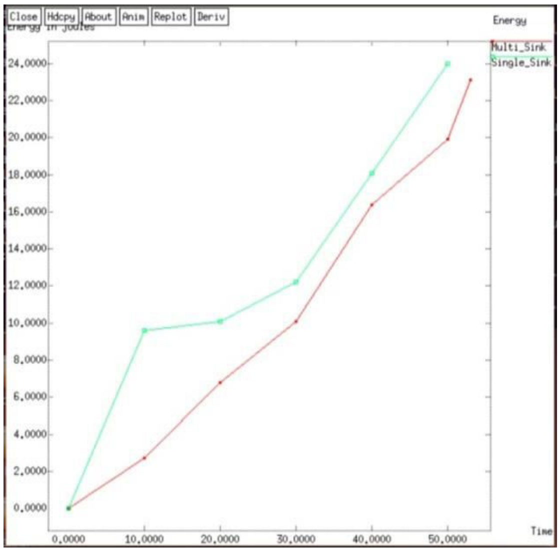

Figure 13: Total energy versus number of nodes.

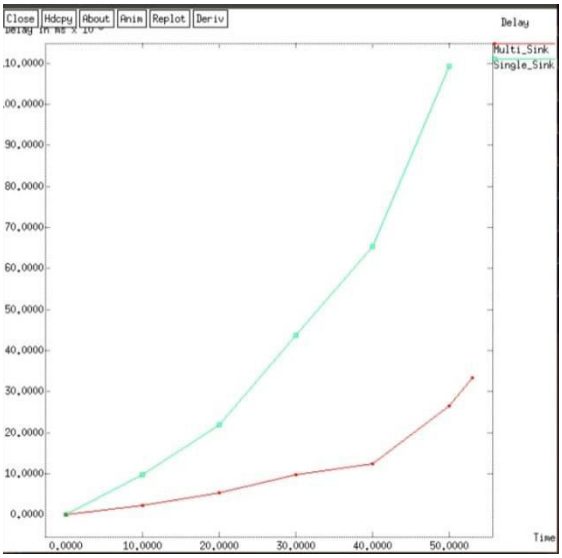

Figure 14: Delay versus number of nodes

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

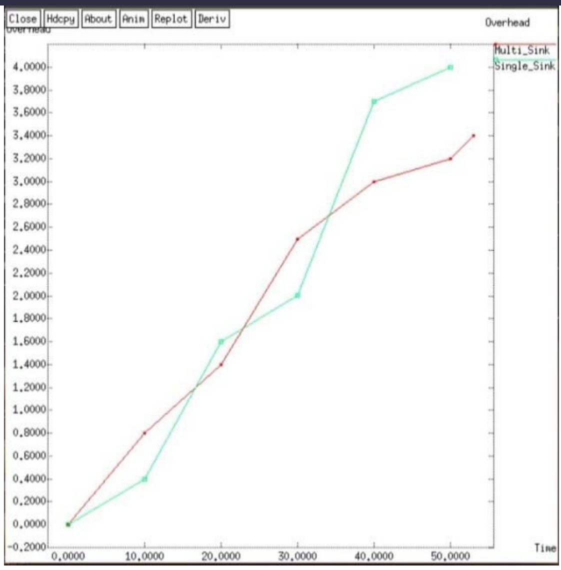

Figure 15: Overhead versus number of nodes.

Figure 14 shows for multi-sink network delay occurred will be less compared to delay occurred in single sink and lastly Figure 15 presents load will be more in single sink compared to multi-sink.

VI. CONCLUSION

The DbCTF scheme is able to meet its goals, reducing the response times of the areas and increasing the average lifetime of nodes. This DbCTF scheme for multiple Base stations was able to maximize the network lifetime by using Data Collection Tree (DCT), Tree-based schemes reduce the communication overhead by minimizing the distance between nodes. The DbCTF strategy for multiple sink shows improved performances than the existing tree-based approaches, in terms of energy consumption, throughput, PDR and overhead. The design of the DbCTF scheme by considering the dynamic (movable) base stations in the heterogeneous environment is left as future work.

REFERENCES

[1] Zhen Hong, Rui Wang, and Xile Li "A Clustering-tree Topology Control Based on the Energy Forecast for Heterogeneous Wireless Sensor Networks". IEEE/CAA journal of automatica sinica, vol. 3, no. 1, January 2016

[2] Chuan Zhu1, Shuai Wu, Guangjie han, Lei Shu,and Hongyi Wu ,"A Tree-Cluster-Based Data-Gathering Algorithm for Industrial WSNs With a Mobile Sink", IEEE Access/Digital ObjectIdentifier10.1109/ACCESS.2015. April 22, 2015.

[3] R.Velmani and B.Kaarthick, "An Efficient Cluster-Tree Based Data Collection Scheme for Large Mobile Wireless Sensor Networks", IEEE SENSORS JOURNAL, VOL. 15, NO. 4, APRIL 2015

[4] Ruitao Xie and Xiaohua Jia, Fellow," Transmission-Efficient Clustering Method for Wireless Sensor Networks Using Compressive Sensing", IEEE transactions on parallel and distributed systems, vol. 25, no. 3, march 2014

.