Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 5, Issue. 12, December 2016, pg.56 – 65

Using Clustering Technique to Improve the Network Lifespan

Arti Bahuguna
M.Tech in Computer Science & Engg., G.B.P. Engineering College Pauri (Uttarakhand)
artibahuguna999@gmail.com

Upendra Bhatt
M.Tech in Computer Science & Engg., G.B.P. Engineering College Pauri (Uttarakhand)
upendra.bhatt12@gmail.com

Shail Ratna Bhatt
CSED HNBGU Srinagar, Garhwal (Uttarakhand)
shailrbhatt@gmail.com

ABSTRACT: The fastest growing technology that would dominate the future world of wireless communication is Wireless Sensor Networks (WSNs).The critical issue in WSNs is energy. In order to improve the lifetime of the network, energy should be used in an efficient manner. Wireless Sensor Network (WSN) consists of small nodes with sensing, computation, and communications capabilities. Sensor node senses the data and sends data to the base station for further processing. These sensor nodes mainly rely upon batteries for energy, which get drained at a quicker rate due to the computation and communication and this decreases the lifetime of the network. In order to solve this problem, various clustering techniques are introduced to improve the lifetime of WSN. The main goal of clustering algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. In our proposed algorithm they are selected in a way that, the distance between them is maximum and energy of the selected CH is more.

Keywords: Cluster Head, Cluster, Wireless Sensor Network.

1. INTRODUCTION

Wireless sensor network (WSN) is constructs by various amounts of node. Sensor node is able to execute sensing information from its area and communicating with each other nodes in the system. Nodes are usually projected to function based on battery lifetime. It is very complicated to renovate batteries of sensor node in the region, nodes are expensive, and energy is constrained for each and every node in WSN. Advances in wireless communication made it possible to develop network consisting of tiny size, little power and multifunctional nodes. A sensor node can sense temperature, pressure, seismic, thermal, video, sound, soil etc. and send information to base station [1]. In densely deployed sensor network, the data sensed by the sensor will be similar and it will generate the same information. If we will transmit this data to the BS then duplicate data will be transferred and great extent of energy will be utilized. Therefore, in order to transmit only unique information, it is advisable to arrange sensor nodes in some group. By arranging sensor nodes in the group, the data will be combined from all the sensor nodes and only compact data will be transmitted to the base station. This will decrease the amount of data that need to be transmitted and it will consume less energy. The process of grouping of sensor nodes is known as clustering. The remaining of this paper is arranged as follows; section 2 reviews the different existing schemes of cluster formation. Section 3 gives the overview of proposed work. Section 4 provides implementation details of proposed scheme and finally section 5 conclude the paper.

2. RELATED WORK

Clustering plays very important role in WSN. In clustering, CH transmits collected data to the BS. Member nodes do not communicate directly with the BS. So, less energy of member nodes is consumed and this reduces the communication overhead. Clustering provides various advantages like scalability, coverage, robustness, simplicity, load balancing and extended network lifetime. In this section, we present a literature survey of existing clustering algorithms for WSN. Various clustering algorithms have been proposed by researchers. LEACH [5] is the basic distributed clustering algorithm. Various algorithms were developed in order to overcome limitations of LEACH. W. Heinzelman et al. [5] proposed first dynamic clustering algorithm. In this algorithm, the CH is not same in each round. Roll of CHs is revolved between members of clusters to distribute the communication cost within the cluster. This protocol helps in reducing amount of data that needs to be transmitted to the BS. There are some limitations of LEACH algorithm like it does not provide optimal number of CHs. Energy parameter is not considered while selecting the CH. There is no even distribution of the CHs. Single hop clusters are formed which will consume more energy because data will be transferred directly from sensor nodes to BS.

In order to overcome limitations of LEACH, new improved clustering algorithm was developed by W. Heinzelman et al. [6]. LEACH-C is an algorithm where CHs are decided by BS. Decision regarding to a sensor node to join which cluster is also made by BS. In setup phase, every node send data related to energy level and latest position to BS. BS decides one threshold level and the nodes that have energy higher than the threshold is selected as CH. After determining CH, BS sends a message containing all the CHs. When node receives this message it compare CH ID with its own ID and if it matches then it is a CH else it is
a cluster member.

R. Randriatsiferana et al [11], proposed algorithm to rectify work of LEACH algorithm. This algorithm changes the metrics for electing a CH and finding the optimal number of CHs. The optimal number of CHs is based on the density of the covering. Nodes with the highest residual energy and the lowest energy variance consumption become cluster-heads. The variance parameter keeps energy consumption dispersion. This approach provides optimal number of clusters and it is energy-efficient.

B. Tang et al [10], proposed a Geographic Energy Aware Routing Centralized Clustering (GEAR-CC) algorithm. Being a centralized clustering algorithm it utilizes the high power of BS to conduct every transmission in WSN. BS formulates the optimal transmitting schemes for all sensor nodes by using global information of topology and energy. The optimization is achieved by making trade-off between energy cost and node's residual power. Random selection of CH is done in order to transfer the data to the BS. So, it solves the hot-spot problem. All the routes are globally optimized because best route is found by BS for each node.

A novel distributed algorithm was developed by O. Younis et al [7]. In Hybrid Energy Efficient Distributed (HEED) communication between cluster heads is done using intermediate nodes and communication between cluster members is done without help of intermediate node. Communication that is done using intermediate nodes is called multi-hop communication and communication that is done without intermediate node is called single-hop. Two factors are considered while selecting CH: remaining energy and cost of communication between cluster members. Preliminary set of CHs are selected using remaining energy of nodes and intracluster communication cost is used to determine magnitude of node or node's distance with its neighbors. When clusters are formed node selects that CH which requires less intracluster communication cost.

A unique centralized clustering algorithm was developed by S. Murguganathan et al [8]. In Base-station Controlled Dynamic Clustering Protocol (BCDCP) base station makes decision about CHs, routing paths and rotation of CHs. The main focus is to create balanced cluster that is cluster with equal number of members and uniform placement of CHs across the network. This algorithm solves the hot spot problem because the nodes that are transmitting data to the BS are changed in every round. Clusters are formed using cluster splitting algorithm [8]. If balanced clustering technique is used to form the cluster then clusters are balanced but it is not be energy efficient. A new algorithm was developed by D. Weigh et al [9] that finds the size of cluster based on how far the base station is. This Energy-Efficient Clustering (EC) algorithm achieves estimation of node lifespan and reduces energy utilization. Each sensor node makes observations and based on this a unique data packet is transmitted to its CH. After collecting all the unique data packets from cluster members, a single summary packet is produced. This summary packet represents the cluster. This procedure is called a single data collection round. Energy is used effectively in this algorithm and solves the hot-spot problem.

There are two phases in every round: setup phase and steady phase. Clusters are created in the setup phase and data are transmitted during steady phase. Advantage of this algorithm is that there is no single CH that remains the CH in next round. Different set of CHs are selected for each round. It is energy efficient then the static method of clustering and direct diffusion [5].

3. PROPOSED WORK

Limitation of LEACH [5] and LEACH-C [6] is that the cluster head selection is done on the basis of the initial energy. Residual energy of the nodes is not taken into consideration. In HEED [7], the nodes near the base station may die earlier because they have more data to transmit to base station and energy consumption is also not equal. In BCDCP [8], performance of the algorithm decrease as the sensor field becomes small. In GEAR-CC [10] clusters are arranged geographically. It doesn’t follow any particular method of cluster formation. Nodes in the network don’t know how to join a cluster or they belong to which cluster or how to choose a best path. In order to overcome limitations of existing we are proposing a new clustering algorithm that improves energy consumptions of nodes. If the energy consumption is managed efficiently than network lifespan will be improved.

In our proposed algorithm, CH is selected using two parameters: distance between nodes and residual energy of nodes. By considering these parameters while selecting CHs the most eligible CHs are selected as CH. In this algorithm, we assume that location of nodes is known in advance. Location of nodes can be found using GPS or any localization techniques like range based techniques or range free techniques. Our proposed algorithm is running at the BS. Pseudo code for proposed algorithm is given below:

Pseudo code: Proposed method for clustering in WSN

```
BEGIN
 INITIALIZE_NETWORK ( )
 Estimate OPT_CH
 WHILE NOT WSN_DEAD ( )
  FOR each node, i, in the network DO
     IF Res_Energy(i) ≥ AVG_NW_Energy
        Prepare ELIGIBLE_CH_LIST
  END FOR
  Select CH from ELIGIBLE_CH_LIST
   CLUSTER_FORMATION ( )
  FOR each CH, i, in the network DO
        FIND NEAREST CH
        i Send Data to NEAREST_CH
  END FOR
   Selects   the   CH   nearest   to   BS,
CH_MIN_DIST_BS
  CH_MIN_DIST_BS Sends Data to BS
  END WHILE
END
```

Pseudo code of proposed algorithm consists of some functions and variables which are described below.

INITIALIZE NETWORK ( ): This function will initialize the network in terms of total number of nodes, area of the network, node id, node type, node energy etc. Node type is used to identify that a sensor nodes is CH or member node.

OPT CH: This is variable which finds the optimal number of CHs based on network parameters. It is calculated from equation:

$K = \sqrt{N}/2*3.14*\sqrt{Energyfree\ space/Energymulti\ path}*db^2_{bs}$ Where K is optimal number of CHs, N is total number of nodes in the network, Energy. free space is energy required in free space, Energy. multi path is the energy required for multi path, M is the area of the network and dbs is the distance from BS [6].

WSN DEAD ( ): This function is used to determine when to terminate the algorithm. This function will check energy of the network and if energy of all the nodes fall to 0 then network is considered to be dead.

Res Energy: This is a variable which stores the residual energy of the node.

AVG NW Energy: This variable stores the average energy of the network after each round of communication. The procedure of selecting CHs, clusters formation and data are transmission to the BS is called one round. When data is transmitted to the BS, one round is said to be complete.

ELIGIBLE CH LIST: This is used to calculate the average energy of the network. Then residual energy of each node is compared with average energy of the network. A list of eligible CH is made by considering those nodes whose energy is greater than average energy. From this list CHs are selected in such a way that, the distance between them is maximum and energy of the selected CH is more. Complexity for finding the CH is O(N2)

CLUSTER FORMATION ( ): This function is used to create clusters. The distance between a node and every CH is calculated and whichever CH is closer to it; will be the CH of a node. It also keeps a sequence of CHs according to the distance from the BS. The farthest located CH will be placed at the top of the list. Complexity for forming the cluster is O(K*N)

NEAREST CH: This variable stores the distance between each CH. Each CH calculates its distance from all other CHs and stores the CH whichever is closer. This is a member variable of a node structure. Node structure contains information about node location, residual energy, node id, distance to the CH, etc. The CH which has maximum distance from BS, starts transmitting data to CH. CH MIN DIST BS: This variable stores the id of CH that is closer to the BS and all the CH will send its data to this CH and this CH will send the data to the BS.

4. IMPLEMENTATION DETAILS

We simulated our proposed algorithm using MATLAB with a network having random distribution of sensor nodes across the field. We selected MATLAB as our simulation tool because many researchers like (8) and (9) are using MATLAB for their simulation work. We have used first order radio model [5] for the energy calculations. Equations (2) and (3) are used for energy calculations. For simulation, we have taken following constant values:

E_{elec} = 50 nJ/bit, E_{amp} = 100 pJ/bit/m2, E_{da} = 5 nJ/bit.

$ETx(k; d) = Eelec *k + Eamp *k * d*d$…………………(2)

Where E_{Tx}(k, d) is the energy required to transmit k bit message to distance d. E_{elec} is the transmitter electronics. E_{amp} is the transmitter amplifier.

$ERx(k) = Eelec *k$ ……………………..(3)

Where E_{Rx}(k) is the energy required to receive k bit message. Network of 400 nodes are shown in Figure 1. Location of BS is at (x=0, y=0) coordinates, and network diameter is 100 m x 100 m. Each node has 4000-bit data packet to transmit and size of control packet is 200 bits. Control packet includes information about energy levels and location of nodes. Figure 2 represents the clusters that are formed when number of nodes are 400 and total number of CHs is 3.

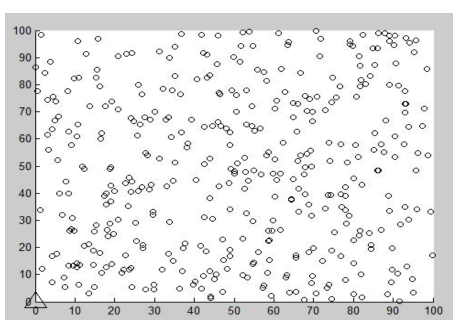

Fig.1 Random Distribution of 400 sensor nodesqsf file converted to simpler tabular form

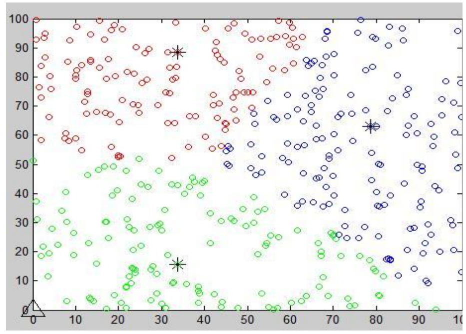

Fig.2 Cluster formation 400 nodes

Figure 3 shows the number of sensor nodes that are dead after 600 rounds with each node initially given 0.2 J of energy. Analysis of proposed algorithm is accomplished using different number of nodes, different value of path loss exponents and different number of rounds. Proposed algorithm is compared with existing two algorithms: LEACH [5] and BCDCP [8]. Performance of the proposed algorithm is measured using total number of dead nodes per round and total residual energy of the network per round.

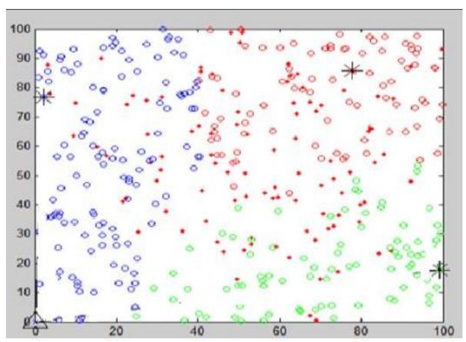

Fig.3 Number of dead node in the network after 600 round

Fig 4 represents the comparison of dead nodes in the network of 400 nodes after 600 rounds. The proposed algorithm has less number of nodes that are dead compared to other two algorithms. This experiment is performed by considering path loss exponent as 2 that is free space.

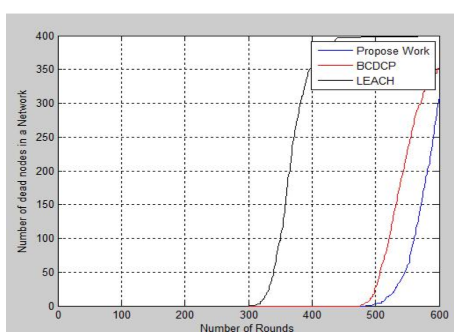

Fig.4 Comparison of No.of dead nodes in network for path loss exponent=2

Fig 5 represents the comparison of residual energy of network after 600 rounds. The proposed algorithm has more residual energy compared to other two algorithms. This experiment is performed by considering path loss exponent as 2 that is free space and total number of nodes is 400.

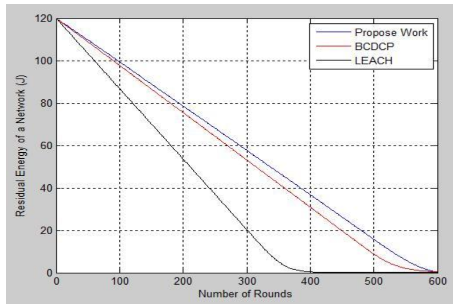

Fig.5 Comparison of residual energy of network for path loss exponent=2

Similarly, Fig 6 and Fig 7 represent the comparison of number of dead nodes and residual energy of network after 600 rounds for the path loss exponent value 4.

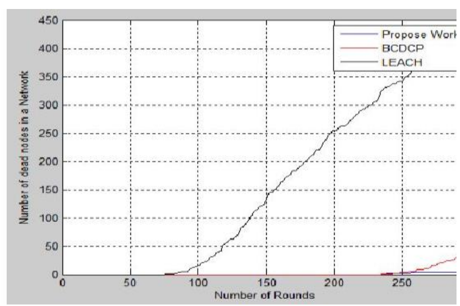

Fig.6 Comparison of Comparison of Number Dead Nodes in the network for path loss exponent=4

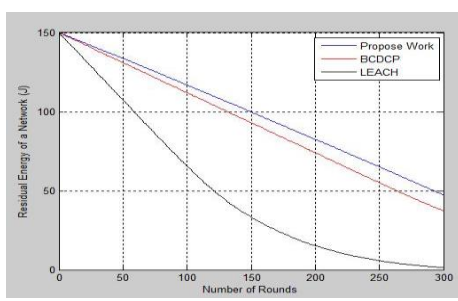

Fig.7 Comparison of Residual Energy of Network for path loss exponent=4

In our proposed algorithm, CHs are not selected only on the basis of energy of the nodes but the distance between CHs is also taken into consideration. For data transmission, CH that is closer to the BS is selected. All other CH sends data to this CH. Performance of our proposed algorithm is improved then existing clustering algorithms because of our two selection criteria, clusters are formed uniformly and that's why energy consumption is also unique. Clusters are distributed equally in the sensing area rather than only on the extreme corners of the sensing area. Our proposed algorithm is more efficient than the LEACH because in our proposed algorithm BS have global knowledge about location and energy level of all the nodes in the network. So that it produces better clustering that requires less energy for data transmission.

Our algorithm improves the performance from BCDCP because in BCDCP they try to balance the cluster. So cluster members have to use more energy to transmit data to the respective CH.

Table 1 represents the value of total number of dead nodes and residual energy of the network for different network density. From this analysis we can say that our proposed algorithm improves performance of network in terms of number of dead nodes and residual energy of the network. If our proposed algorithm is able to save energy then it can prolongs the network lifetime dramatically.

Nodes in the n/w	Number of dead Nodes			Residual Energy (J)		
	LEACH[5]	BCDCP[8]	Proposed	LEACH[5]	BCDCP[8]	Proposed
100	89	75	15	0.0646	0.1597	0.4237
200	192	170	112	0.0587	0.4277	0.2294
300	298	277	117	0.0258	0.2493	1.135
400	399	353	309	0.0047	0.3917	0.3843
500	499	498	112	0.0077	0.0083	5.7280

TABLE I. PERFORMANCE ANALYSIS USING PATH LOSS EXPONENT = 2

5. CONCLUSION

In various applications of WSN the energy recharge is not possible. By applying clustering in WSN, nodes have less data to transform to the base station and cluster head applies data fusion. By clustering we can optimize the use of energy and if the energy is conserved then network lifetime can also be increased. In our proposed algorithm, CH is selected using distance between nodes and energy of the nodes. Our proposed approach uses the cluster head that is closer to the BS to transmit the collected data to the base station. Selection of cluster heads varies in every round of communication. We have evaluated our work for different network density and path loss exponent.

REFERENCES

[1] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y.,Cayirci, E. (2002). Wireless sensor networks: a survey. Computer networks, 38(4), 393- 422.

[2] Al-Karaki, J. N., Kamal, A. E. (2004). Routing techniques in wireless sensor networks: a survey. Wireless communication, IEEE, 11(6), 6-28.

[3] Yick, J., Mukherjee, B., Ghosal, D. (2008). Wireless sensor network survey. Computer networks, 52(12), 2292-2330.

[4] Lee, M. G., Lee, S. (2007, May). Data dissemination for wireless sensor network. In Object and Component-Oriented Real-Time Distributed Computing , 2007. ISORC'07. 10th IEEE International Symposium on (pp. 172-180). IEEE. vHeinzelman, W. R., Chandrakasan, A., Balakr- ishnan, H. (2000, January). Energy-efficient communication protocol for wireless microsensor networks. In System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on (pp. 10-pp). IEEE.

[5] Heinzelman, W. B., Chandrakasan, A. P., Balakrishnan, H. (2002). An application specific protocol architecture for wireless microsensor networks. Wireless Communications, IEEE Transactions on, 1(4), 660- 670.

[6] Younis, O., Fahmy, S. (2004). HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. Mobile Computing, IEEE Transactions on , 3(4), 366-379.

[7] Muruganathan, S. D., Ma, D. C., Bhasin, R. I., Fapojuwo, A. (2005). A centralized energy-efficient routing protocol for wireless sensor networks. Communications Magazine, IEEE, 43(3), S8-13.

[8] Wei, D., Jin, Y., Vural, S., Moessner, K., Tafazolli, R. (2011). An energy-efficient clustering solution for wireless sensor networks. Wireless Communications, IEEE Transactions on, 10(11), 3973-3983.

[9] Tang, B., Wang, D., Zhang, H. (2013, October). A Centralized Cluster- ing Geographic Energy Aware Routing for Wireless Sensor Networks. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics (pp. 1-6). IEEE Computer Society.

[10] Randriatsiferana, R. S., Alicalapa, F., Lorion, R., Mohammed, A. M. (2013, September). A clustering algorithm based on energy variance and coverage density in centralized hierarchical Wireless Sensor Networks. In AFRICON, 2013 (pp. 1-5). IEEE.

