ENHANCED NETWORK LIFETIME USING HYBRID K-MEANS CLUSTERING IN WIRELESS SENSOR NETWORKS

Senthamarai.K, Assistant Professor, Computer Science and Engineering, Sri Ramakrishna Institute of Technology, Coimbatore 641010, Email: senthamarai.srit@gmail.com

ABSTRACT:

Wireless sensor network is a network of energy constrained nodes with the capability of sensing and communication covering a large area. Although WSNs have significant advancements in many areas; maximizing the lifetime of the whole network remains a major obstruction.In thisstudy, an energy efficient clustering protocol based on hybrid K-means algorithm has been proposedfor WSN where midpoint algorithm is used to improve initial centroid selection procedure.Midpoint method has been applied forinitial CHs selection, instead of choosing initial CHs randomly. Itobtains balanced cluster where CHs are uniformly distributed andeach cluster contains an almost equal number of sensor nodes. Asa result, the load of the CHs becomes balanced, which ultimatelyprolong the network lifetime.Simulation result alsodemonstrates that the proposed approach can reduce the energy consumption at most 60% compared to LEACH-Bprotocol.

Keywords: Lifetime, Midpoint algorithm, Hybrid K-Means, centroid selection.

INTRODUCTION:

A sensor node generally consists of sensors, actuators, memory, a processor and they communicate through a wireless medium.WSNs are deployed to carry out various applications, such as environmental monitoring [1], wildlife habitat monitoring [2], acoustic monitoring [3], and battlefield surveillance [4]. But these applications raise a few challenges like improving the network lifetime by making the sensor nodes functional for a longer period of time. Figure1 represents a typical wireless sensor network. The sensor nodes (SN) send the data packet to the respective cluster heads (CH) which further sends data to the sink or the base station. The nodes can vary from hundreds to thousands.

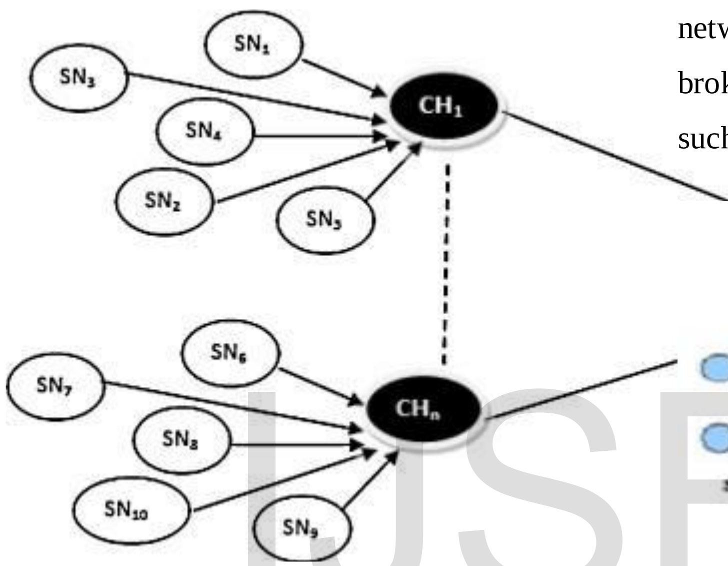

Fig.1. Typical Wireless sensor networks [5]

Sensor nodes or motes in WSNs are small sized and are capable of sensing, gathering and processing data while communicating with other connected nodes in the network, via radio frequency (RF) channel. Architecture of the wireless sensor network [6], Consists of the following components.

(a) Sensor node: a sensor node is the core component of the wireless sensor network because it can perform multiple roles in a network, such as sensing, data processing, data storage and routing.

(b) Clusters: Clusters are the organization unit for WSNs. The dense nature of these networks requires the need for them to be broken down into clusters to simplify tasks such a communication.

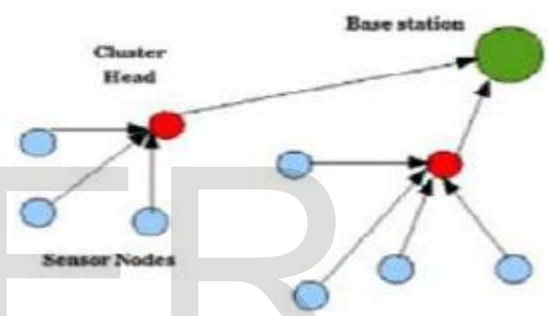

Fig.2. Basic Cluster Structure

(c) Cluster Heads: cluster heads are the organization leader of a cluster. They often are required to organize activities in the cluster. These tasks include but are not limited to data-aggregation and organizing the communication schedule of a cluster.

(d) Base Station: the base station is at the upper level of hierarchical WSN. It provides the communication link between the sensor network and the end user.

(e) End-User: The data in the sensor network can be used for a wide range of applications [3] .Therefore for a particular application may make use of the network data over the internet, using a PDA, or even a desktop computer. All the processing is done at the end user when data from various nodes are collected.

LEACH-B (LEACH-Balanced), at eachround, after first selection of cluster head according to LEACHprotocol, a second selection is introduced to modify the numberof cluster head in consideration of node's residual energy. As aresult the number of cluster head is constant and near optimalper round.But, the constant cluster head will not satisfy all the time period.

In this paper, the proposed EECPHK-means protocol improves the initialcentroid selection procedure of K-means using midpoint algorithm[7] which produces balanced cluster compared to K-means. Italso optimizes CH selection method by considering residualenergy as a parameter of CH selection in addition to Euclideandistance used in basic K-means algorithm.

The rest of the paper is organized as follows: Section 2 describes the proposed EECPHK-means protocol. Section 3 shows the experimental results and comparison. Section 4 concludes the paper.

II. PROPOSED SYSTEM:

Network model

The following assumptions are made regarding the network model:

(i) The sensor nodes and BS are all static after deployment.

(ii) There is only one BS far from the sensing region.

(iii) Sensors are homogeneous having the same initial energy.

(iv) After deployment BS knows the geographical information of allsensor nodes.

(v) Data aggregation happens in each of the CH. The CH ultimatelysends aggregated data to the BS.

(vi) CHs follow single-hop or multi-hop communicationsdepending on the distance from the BS.

Algorithm 1: The Hybrid K-means clustering algorithm

1)**repeat**

(2) find two objects(clusters) with closest distanceamong all, and cluster them together.

(3) the value of attributes of new cluster are theaverage of attributes of two old objects(clusters);

(4)**until**the percentage of hierarchical clustersrequested by user is done;

(5) calculate average attribute values of members ofclusters that generate from step (1) to (4) as initialcluster centroids;

(6)**repeat**

(7) **for**all objects

(8) **if**the object already appeared in step(2)

thenthe object remain in original cluster;

(9) **else**

Calculate distances between the object andexisted clusters

(10) **if**the shortest distance lower thanthreshold

(11) **then** the object are assigned to theclosest cluster

(12) **else**

the object belongs to minor group

(13) **end for loop**

(14) update the centroid attribute value;

(15)**until**no member changes belonging cluster;

In algorithm 1 describes the hybrid K-Means clustering algorithm.

The working principle of the proposed EECPK-means protocol isdivided into three phases as follows:

Phase 1: Initial CH selection.

Phase 2: Balanced cluster formation.

Phase 3: Data communication.

These three phases are described in Algorithm 2, Algorithms 3and 4 (see Figs. 2 and 3), respectively.

Algorithm 2: Midpoint algorithm for initial CH selection

Input:

D = set of n data points.

kopt = optimum number of desired clusters

Output:

kopt number of initial centroids.

Steps:

1: Compute the distance from origin for each data point.

2: Sort the distances obtained in step 1. Sort the data points inaccordance with the distances.

3: Partition the sorted data points into kopt equal sets.

4: In each set, take the middle point as the initial centroid.

Algorithm 2 describes the midpoint algorithm [9] which has beenused for initial CH selection assuming that the data points containonly positive values. Here the centroid of a cluster is a virtual nodelocating at the centre position of the cluster. To maintain theconnectivity of the network, residual energy of the CH is checkedevery round. If the energy of the CH is smaller than the thresholdenergy, the node having the next ID number is selected as a newCH as in [10]. The newly elected CH informs other nodes aboutthe change of the CH.

Algorithm 3: Balanced Cluster Formation

Input

D= Set of n data items

k_{opt} = Optimum number of desired clusters

$E_{threshold}$ =Threshold energy

Output:

A set of k_{opt} clusters

Steps:

1. Apply Midpoint method presented in Algorithm 2 to choose k_{opt} out of D sensor nodes as initial cluster heads;
2. Repeat
3. Each of the remaining nodes decides to join its nearest CH according to the Euclidean distance.
4. Centroid of each cluster is calculated as

$$Centroid(X,Y) = \left[\frac{1}{S}\sum_{i=1}^{s} x_i, \frac{1}{S}\sum_{i=1}^{s} y_i\right]$$

5. After cluster formation, based on the distance from the centroid, an ID number isallotted to each node of a cluster, assigning smaller number to the closer one.
6. **For** all selected cluster Heads
7. **If** (Residual energy of cluster head $\geq E_{threshold}$)
8. **Then**
9. The node will remain as cluster head
10. **Else**
11. Check ID number of all the sensor nodes in that cluster

12. The node in the next order of ID number is selected as a new CH.
13. End if
14. End for
15. The newly elected CHs inform other nodes about the CHs change.
16. Until the cluster heads are not changed any more.

Algorithm 3 describes the balanced cluster formationphase. In thisapproach, an estimation of thresholdresidual energy is given, which was not addressed in [10]. Here thethreshold energy is the amount of energy required to receive,aggregate and transmit the average number of sensor nodes in thecluster.

SIMULATION RESULT

For simulating the proposed EECPHK-means protocol, MATLAB R2015a and C language have been used. Considering dBS = 100, the numberof desired CH = 4 is obtained and considering dBS = 85, the number ofdesired CH = 5 is obtained, having 100 sensor nodes in 100 × 100 m2 sensingregion. Simulations have been performed for both 4-cluster and5-cluster networks. Theproposed EECPHK-means protocol hasbeen compared with LEACH-B protocol [11] used inWSN with respect to cluster formation.

Parameters for Simulation

Similar random 100-node networks and radiomodels are used with LEACH. This paper set the followingparameters for Simulation:

1) Energy required in sending or receiving 1bit:E_{elec}=50nJ/bit

2) The amount of data sent by nodes each time: k = 200bit.

3) The initial energy ofevery node: E = 0.5J

4) Compress ratio during data fusion: r = 0.7

5) Energy consumed in every bit data fusion: EA=50pJ/bit

6) Area: 100*100

7) The location of Sink: (5, 5)

8) The percentage of cluster head: p = 0.05

9) The number nodes: n = 100.

Figures and Analysis

The network performance is measured by its life time.Figure 3 shows the system lifetime using the EECPHK-means protocoland LEACH-B algorithm.As shown in the figure, nodes in LEACH-B protocolbegin to die after about 800 rounds, and almost all

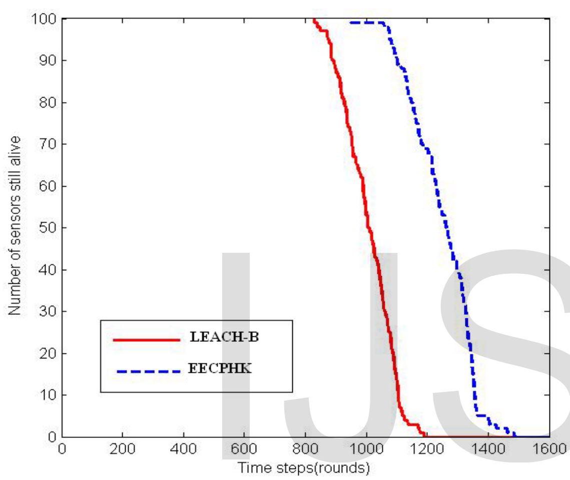

Fig 3: System lifetime

nodesdie when it comes to 1200 rounds. While in EECPHK-means,nodes begin to die after about 1000 rounds, and all nodesdie after 1500 rounds. Therefore, the lifetime of EECPHK-meanshas obvious longer than the LEACH-B protocol whichproves that the improved algorithm has extended thesystem lifetime effectively.

III. CONCLUSION

In this paper, with the research of the LEACH-B protocol,a novel protocol called EECPHK-means is put forward. Thisenhanced algorithm has overcome the shortcoming of theoriginal protocol by taking the node's residual energy intoconsideration and keeping the constant and near optimalnumber of cluster heads at each round.

The simulation result shows that EECPHK-means provides betterenergy efficiency and longer network life span than LEACH-B.

IV. REFERENCES

[1] T. D. R¨aty, "Survey on contemporary remote surveillance systems for public safety", IEEE Trans. Syst., Man, Cybern., Part C, vol. 40, no. 5, pp. 493–515, Sep. 2010.

[2] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Aderson, "Wireless sensor networks for habitat monitoring". In Proc. ACM International Workshop on Wireless Sensor Networks and Applications, (2002), pp. 88–97.

[3] H. Wang (2006). “Wireless Sensor Networks for Acoustic Monitoring”, Doctoral Thesis, (2006), SBN: 978-0-542-79693-7.

[4] K. R¨omer, F.Mattern, and E. Zurich, “The design space of wireless sensor networks” ,IEEE Wireless Commun., vol. 11, no. 6, pp. 54–61, Dec. 2004.

[5] Murad A. Rassam, AnazidaZainal, MohdAizainiMaarof, “An adaptive and efficient dimension reduction model for multivariate wireless sensor networks applications”, Journal of Applied Soft Computing, Vol 13 issue 4, April 2013, pp. 1978-1996

[6] Ankita Joshi, Lakshmi Priya M, “A Survey of Hierarchical Routing Protocols in Wireless Sensor Network” MES Journal of Technology and Management May 17, 2011.

[7] Aggarwal, N., Aggarwal, K.A.: ‘A mid–point based k-mean clustering algorithm for data mining’, Int. J. Comput. Sci. Eng., 2012, 4, (6), pp. 1174–1180

[8] M. Tong and M. Tang, "LEACH-B: An Improved LEACH Protocol for Wireless Sensor Network," 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), Chengdu, 2010, pp. 1-4. doi: 10.1109/WICOM.2010.5601113

[9] Aggarwal, N., Aggarwal, K.A.: ‘A mid–point based k-mean clustering algorithm for data mining’, Int. J. Comput. Sci. Eng., 2012, 4, (6), pp. 1174–1180

[10] Park, G.Y., Kim, H., Jeong, H.W., et al.: ‘A novel cluster head selection method based on K-means algorithm for energy efficient wireless sensor network’. Proc. of IEEE 27th Int. Conf. on Advanced Information Networking and Applications Workshops, 2013, pp. 910–915.

[11] M. Tong and M. Tang, "LEACH-B: An Improved LEACH Protocol for Wireless Sensor Network," 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), Chengdu, 2010, pp. 1-4. doi: 10.1109/WICOM.2010.5601113.