

Optimal Clustering Technique Using K-means Algorithm in Wireless Sensor Networks

Yashodha K Bilagi, Chethana N.S.

M.Tech, Senior Scale Lecturer

¹M.Tech Digital Electronics Dept. of ECE SSIT, Tumakuru 572105, Karnataka, India

²Senior Scale Lecturer Dept. of E&C Government Polytechnic, Hiriya 577599, Karnataka, India

ABSTRACT

The advances in microelectromechanical systems (MEMS) technology, wireless communication and digital electronics have given rise to the development of low cost, low power multifunctional sensor nodes which are in smaller size and capable of short distance communication. Most of the times, the sensor nodes have to operate autonomously in an unattended environment where the recharging/replacement of the battery energy is impossible. Therefore, it is very essential to utilize the available sensor node battery energy in an efficient manner.

In this direction, the proposed work considers grouping of sensor nodes into clusters, which effectively increases the scalability & reduces power consumption. Presently, an efficient k-medoids algorithm is used in order to bring about optimal number of clusters which results in improved load balancing & increased lifetime of the network. The proposed work also considers the basic issue of identifying the mobile node location, using samples of RSSI measurement.

Keywords: WSN, RSSI, MAC, clusters, k-medoids.

1. INTRODUCTION

Wireless Sensor Network can be defined as a special class of ad hoc wireless network that can be used to provide a wireless communication infrastructure that allows us to sense, observe and react to the events & phenomena in the natural environment.

The WSN architecture comprises of sensor nodes which are scattered in sensor field & each of the nodes have the capability to collect the data & route the data back to the sink (base station). Data are routed back to the sink (base station) by a multihop infrastructures architecture through the sink. The sink (base station) may communicate with the task manager node via internet or satellite. WSN can be used in many applications[1].

There is no tethered power supply is available in the

network. Some form of batteries are necessary to provide energy. Sometimes, some form of recharging by obtaining energy from the environment is available as well (eg: solar cell). But, in some environment where human operating is not possible, in such environment, there is no way to replace the batteries for communication. Hence, power will become crucial part in WSN. Thus, we can reduce the power consumption by some of the important methods like routing, localization and MAC protocol. Among these three protocols, MAC protocol is more helpful as it consumes less power & is more efficient. MAC protocol controls the communication module of WSN node[3].

One of the advantages of wireless sensors networks (WSNs) is their ability to operate in harsh environments in which human monitoring schemes are risky, inefficient and sometimes infeasible. Thus, sensors are expected to be scattered randomly in the area of interest by a relatively uncontrolled means. Grouping of sensor nodes into clusters has been widely pursued by the research community to achieve the network scalability objective [4].

Wireless Sensor Network are tremendously being used in different environment to perform various tasks such as target tracking, search, disaster relief & number of tasks in smart environment. In such applications, node localization is one of system parameter. Hence, it is one of the challenge in WSN. There are different methods used in localization to estimate the distance between two nodes: RSSI, TOA, TDOA & AOA. RSSI is more preferred method as it gives accurate distance measure of the nodes[7].

2. RELATED WORK & CONTRIBUTIONS

I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci (2003) proposed a survey on Wireless sensor networks. Recent advances in wireless communications, digital electronics and microelectromechanical systems (MEMS) technology have enabled the development of low-cost, low-power, multifunctional sensor nodes that are smaller in size and communicate in short distances. These sensor nodes that consist of sensing, data processing, and

communicating components, gives the idea of sensor networks based on collaborative effort of a large number of nodes.

Another special feature of sensor networks is the cooperative effort of sensor nodes. Instead of sending the raw data to the sensor nodes which are responsible for the fusion, sensor nodes with their processing abilities will carry out simple computations and transmit only the required and partially processed data.

The above described features ensure a wide range of applications for sensor networks. Thus this paper describes the concept of sensor [1].

B. Mamalis, D. Gavalas, C. Konstantopoulos and G. Pantziou (2009) proposed a paper on Clustering in WSN. The use of WSN has grown considerably by pointing out the crucial need for scalable & energy-efficient routing protocols in large-scale environment. This paper reviews about different clustering algorithms used in the network[5].

LEACH	EEHC	HEED
Low Energy Adaptive Clustering Hierarchy	Energy Efficient Hierarchical Clustering	Hybrid Energy Efficient Distributed Clustering
It is an hierarchical, probabilistic, distributed, one-hop protocol	It is distributed, k-hop hierarchical clustering algorithm	HEED is a distributed, hierarchical, clustering scheme
All nodes have chance to become CH in order to balance the energy by each node.	Each node is elected as a CH with probability "p" & announces about its election to the neighboring nodes within "k"-hop range.	Only those sensors that have a high residual energy are expected to become CH nodes.
Drawback: As the decision on CH election & rotation is probabilistic, there are chances that a node with a very low power gets selected as CH and the elected CHs will concentrate on one part of network.	Drawback: The energy consumption for network operation depends on the parameters p and k of the algorithm.	Drawback: A knowledge of the entire network is needed to determine the intracluster communication cost and configuration of those parameters might be difficult in practical world.

A. P. Reynolds, G. Richards, and V. J. Rayward-Smith (2001) proposed a paper on The Application of K-medoids and PAM to the clustering of rules. This paper also describes about K-means algorithm in which each cluster is centered about a point called as centroids, where the centroid's coordinates are the mean of the coordinates of the objects in the cluster. But the disadvantage of k-means is to calculate the distances of the centroids for each object at each iteration. Hence, K-medoids algorithm was introduced. Rather than calculate the mean of the items in each cluster, a representative item, or medoid, is chosen for each cluster for every iteration.

There are two advantages for using existing rules as the centres of the clusters. Firstly, a medoid rule serves to describe the cluster. Secondly, there is no need for repeated

calculation of distances at every iteration, since the k-medoids algorithm can simply look up distances from a distance matrix,

The k-medoids algorithm can be briefed as:

1. Choose k objects at random to be the initial cluster medoids.
2. Assign each object to the cluster associated with the closest medoid.
3. Recalculate the positions of the k medoids.
4. Repeat Steps 2 and 3 until the medoids become fixed.

Thus, this algorithm has excellent feature that it requires the distance between every pairs of objects only once and uses this distance at every iterative step and it takes less computation time. The above algorithm runs just like K-means clustering and so this will be called as 'K-means-like' algorithm[6].

3. SYSTEM DESIGN

The flowchart of the system design is shown in figure 3.1.

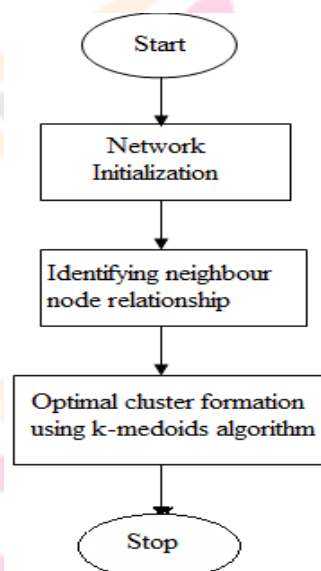


Figure 3.1: The flowchart of the system design

Network Initialization: Initially, a wireless sensor network with a size of about 30-50 nodes is placed randomly on the grid by configuring various network parameters. Nodes are given some mobility to implement mobile network scenario. **Neighbour Node relationship:** By initiating the beacon signals from base station and collecting the reply from the sensor nodes, the Received Signal Strength information is maintained in the base station. The neighbour node identification is accomplished using the collected RSSI information.

Optimal cluster formation using k-medoids algorithm: Based on the computed positions of the nodes, the cluster formation is brought about with few cluster members

coordinated by cluster heads using an algorithm called k-medoids. The shape of clusters changes every time as the nodes are mobile in nature. Each cluster will be given different colors.

4. SYSTEM IMPLEMENTATION

The implementation of algorithm was carried out in MATLAB programming Language. MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment. MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces. The implementation results are as follows:

4.1 Network initialization

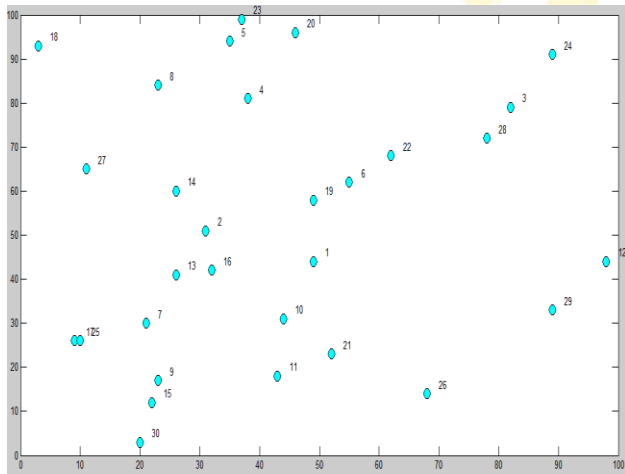


Figure 4.1: Snapshot showing network initialization

The figure 4.1 shows that the nodes are deployed randomly in the network. Each node will be given with some parameters like:

Initial energy = 100J

Communication range = 150m

Communication model = Energy model

Area = 100 x 100m

4.2 Cluster formation

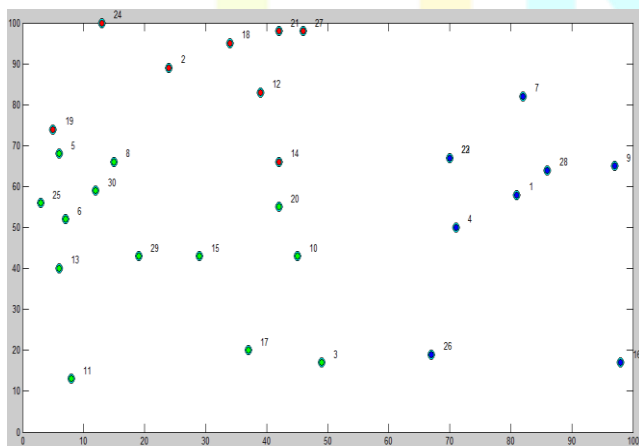


Figure 4.2: Snapshot showing cluster formation

The figure 4.2 shows that each cluster in the network will be

given different colors. These clusters are formed using k-medoids algorithm.

4.3 Identify cluster head & base station

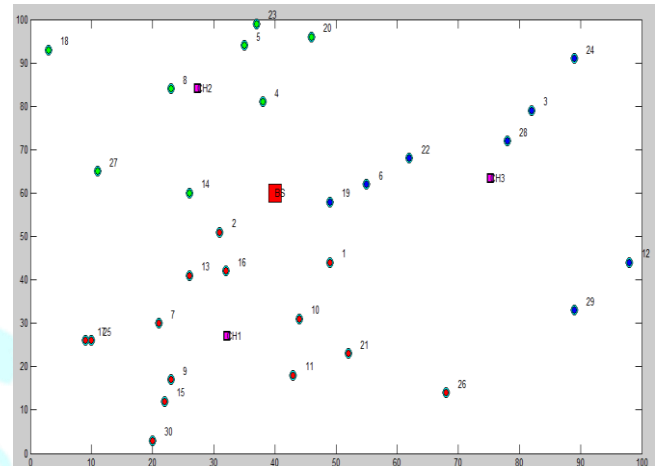


Figure 4.3: Snapshot showing identification of cluster head & base station.

The figure 4.3 shows that for each cluster, a leader called as cluster head will be elected by the nodes based on its energy & location.

5. CONCLUSION

As localization and power control are major issues in the WSN, this paper provide an approach to control and optimize the communication power in the mobile sensor network. The position of the node is estimated using RSSI measurement which can be used in a statistical method called k-medoids algorithm. This results in the formation of optimal number of clusters which is beneficial for scalable & load balancing network.

The MATLAB code is used for simulation of the sensor network as discussed in figures 4.1, 4.2 & 4.3.

Further, the k-medoids method of clustering will be designed & implemented to achieve an improved results with respect to energy consumption, latency & throughput.

REFERENCES

- [1] I. Akyildiz, W. Su, Sankarasubramaniam, and E.Cayirci, "Wireless Sensor Networks: a survey", Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. Received 12 December 2001; accepted 20 December 2001.
- [2] Surendra Bilouhan, Prof. Roopam Gupta, "Optimization of Power Consumption in Wireless Sensor Networks", International Journal of Scientific & Engineering Research Volume 2, issue 5, may-2011, ISSN 2229-5518.
- [3] S. Swapna Kumar, M. Nanda Kumar, V. S. Sheeba, K. R. Kashwan, "Power Management of Hybrid Scheduling Routing in Cluster Based Wireless Sensor Networks",

Journal of Information & Computational Science 9: 6(2012)
1555-1575. Available at <http://www.joics.com>.

[4] Ameer Ahmed Abbasi, Mohamed Younis” A survey on clustering algorithms for wireless sensor networks” Computer Communications 30 (2007) 2826–2841. 0140-3664/\$ - see front matter _ 2007 Published by Elsevier B.V. doi:10.1016/j.comcom.2007.05.024.

[5] Basilis Mamalis, Damianos Gavalas, Charalampos Konstantopoulos and Grammati Pantziou, “Clustering in Wireless Sensor Networks”, Zhang/RFID and Sensor Networks AU7777_C012 Page Proof Page 323 2009-6-24.

[6] A. P. Reynolds, G. Richards, and V. J. Rayward-Smith, “The application of K-medoids and PAM to the Clustering of Rules”. School of Computing Sciences, University of East Anglia, Norwich.

[7] Amitangshu Pal, “Localization Algorithms in Wireless Sensor Networks with the current approaches and future challenges”, Department of Electrical and Computer Engineering, The University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, North Carolina 28223-0001 apal@uncc.edu Tel: 1-980-229-3383.

