Performance Analysis of K-mean Clustering Map for Different Nodes

Subham Bhawsar, Aastha Hajari

Embedded System and VLSI Design, Department of Electronics and Communication Engineering
Shiv Kumar Singh Institute of Technology & Science (SKSITS), Indore, Madhya Pradesh, India

ABSTRACT

WSNs is a group of inexpensive and autonomous sensor nodes interconnected and work together to monitor various environmental conditions such as humidity, temperature, pressure, and other climate changes. Wireless sensor nodes are randomly installed and communicate themselves through the wireless communication medium. In this paper we compared the *simulation results of the K-mean techniques for different numbers of nodes like 50, 100 and 150.*

Keyword: K-mean, Nodes, Clusters.

I. INTRODUCTION

The Wireless sensor networks (WSN) are highly distributed networks of small, lightweight nodes and deployed in large numbers to monitor the environment parameters or system by the measurement of physical parameters such as temperature, pressure, or relative humidity [2]. Each node of the network consists of three subsystems: the sensor subsystem which senses the environment, the processing subsystem which performs local computation on the sensed data, and the communication subsystem which is responsible for message exchange with neighbour sensor nodes. While individual sensors have limited sensing region, processing power, and energy, networking a large numbers of sensors gives rise to robust, reliable, and accurate sensor network covering a wider region [2]. Wireless Sensor Networks (WSNs) typically consist of a large number of low-cost, low-power and multifunctional wireless sensor nodes. Nodes are equipped with sensing, communication and computation capabilities, where they communicate via a wireless medium and work collaboratively to accomplish a common task.

To get more efficient and effective result of K-mean algorithm there have been a lot of research happened in previous day. All researchers worked on different view and with different idea. Krishna and Murty[4] proposed the genetic K-means(GKA) algorithm which integrate a genetic algorithm with K-means in order to achieve a global search and fast convergence.

Components of sensor node

Wireless sensor networks (WSNs) have gained world-wide consideration in recent years, particularly with the proliferation in Micro-Electro-Mechanical systems (MEMS) technology which has facilitated the development of intelligent sensors. Sensor node as shown in Figure 1 may also have additional application dependent components such as a location finding system, power generator and mobilize [1]. The analog signals produced by the sensors based on the observed phenomenon are transformed to digital signals by ADC, and then fed into the processing unit. The processing unit, which is generally associated with a small storage unit, manages the events that make the sensor node collaborate with other nodes to carry out the assigned sensing tasks. A transceiver unit connects the node to the network [1]. The main job of sensor node in a sensor field is to detect the events, perform quick local data processing, and broadcast the data. Energy consumption is a key issue in wireless sensor networks (WSNs). In a sensor node, energy is consumed by the power supply, the sensor, the computation unit and the broadcasting unit. The wireless sensor node, being a microelectronics device, can only be equipped with a limited power source ($\leq$0.5 Ah, 1.2 V) [1].
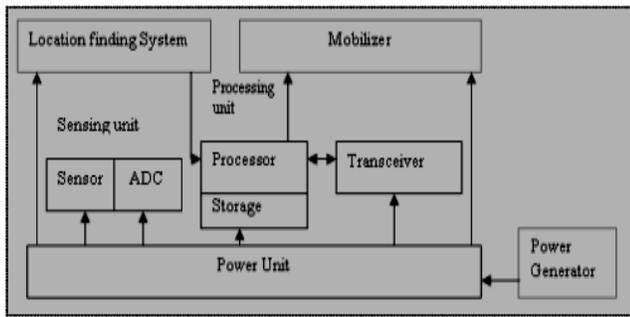
Fig 1: Components of a sensor node [1]

Density based Clustering algorithms: Data objects are categorized into core points, border points and noise points. All the core points are connected together based on the densities to form cluster. Arbitrary shaped clusters are formed by various clustering algorithms such as DBSCAN, OPTICS, DBCLASD, GDBSCAN, DENCLU and SUBCLU [3].
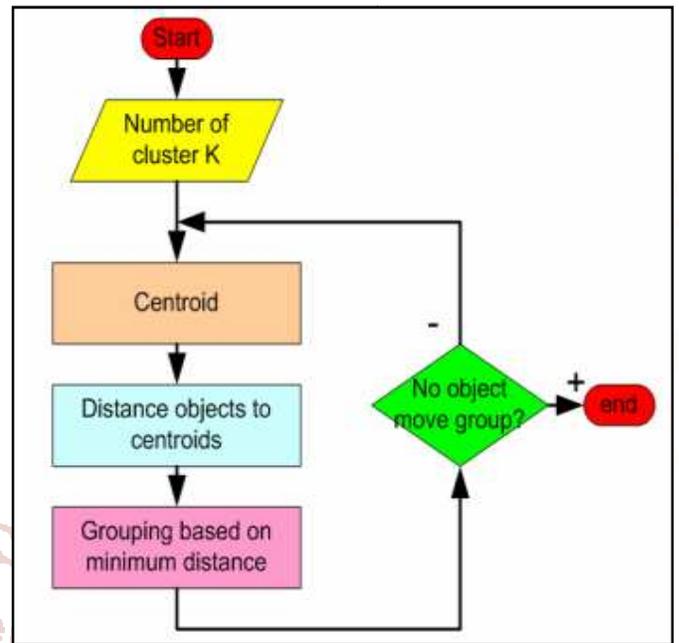
II. K-MEANS CLUSTERING

There are many algorithms for partition clustering category, such as kmeans clustering (MacQueen 1967), k-medoid clustering, genetic k-means algorithm (GKA), Self-Organizing Map (SOM) and also graph-theoretical methods (CLICK, CAST) [6]. k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem, in the procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed apriority. There are main idea is to define k centre's, one for each cluster. These centers should be placed in a cunning way because of different location causes different result. So results, the better choice is to place them as much as possible far away from each other. For the next step is to take each point belonging to a given data set and associate it to the nearest centre. When no point is pending, the first step is completed and an early group age is done.

After we have these k new cancroids, a new binding has to be done between the same data set points and the nearest new centre. A loop has been generated. As a result of this loop we may notice that the k canters change their location step by step until no more changes are done or in other words canters do not move any more.

A. The process of k-means algorithm:
This part briefly describes the standard k-means algorithm. K-means is a typical clustering algorithm in data mining and which is widely used for clustering large set of data's.

Fig. 2: K-Means Algorithm [5, 6]

A.1 K-mean Algorithm Process
1. Select K points as initial centroid.
2. Repeat.
3. Form k clusters by assigning all points to the closest centroid.
4. Recomputed the centroid of each cluster Until the centroid do not change.

III. Result Analysis of K-Means Technique
k-means techniques is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. We consider five access points of K-means technique for 150, 50 and 100 Nodes as shown in Figure 3, Figure 4, and Figure 5, respectively.
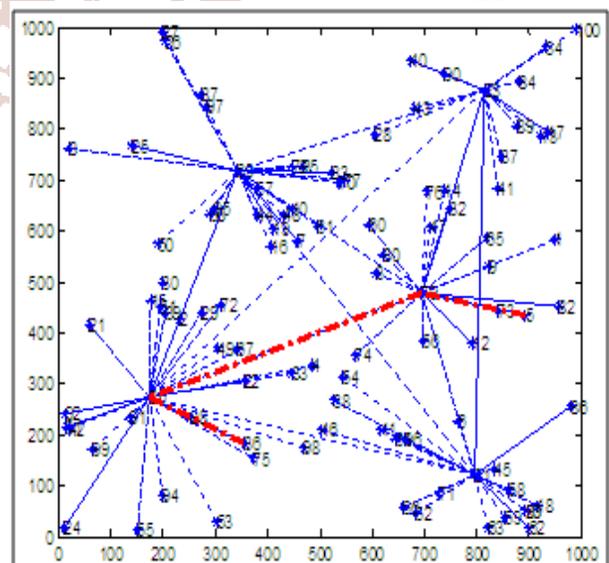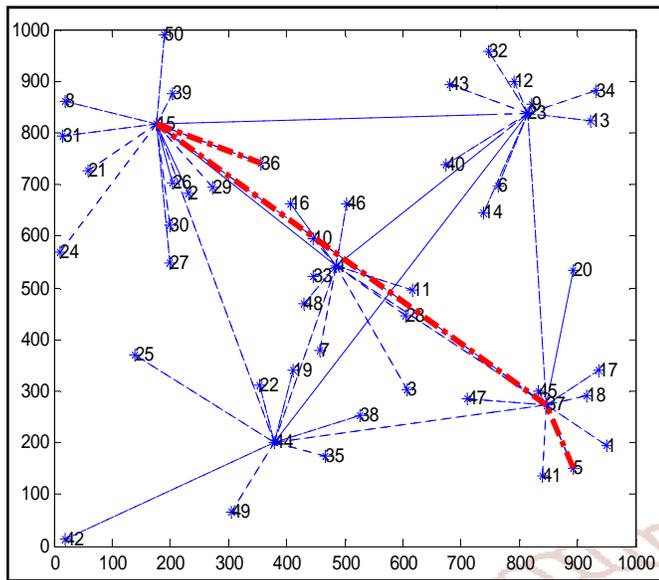

Fig. 3: K-means clustering map for 150 Nodes

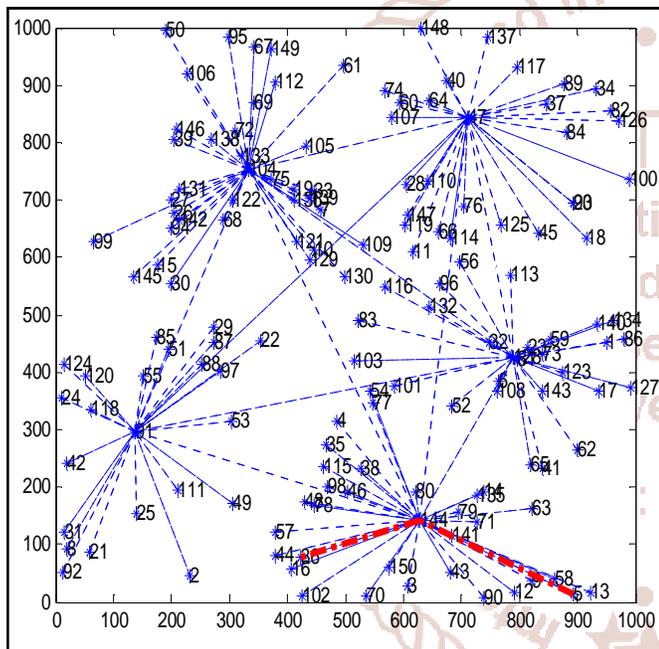**Fig. 4: K-means clustering map for 50 Nodes**

F**ig. 5: K-means clustering map for 150 Nodes**

IV. CONCLUSION

The K-means algorithm is a popular clustering algorithm. Distance measure plays a important rule on the performance of this algorithm. Also, we will try to extend future this work for another partition clustering algorithms like K-Medoids, CLARA and CLARANS.

REFERENCES

1. Vibhav Kumar Sachan, Syed Akhtar Imam "Energy-efficient Communication Methods in Wireless Sensor Networks: A Critical Review" International Journal of Computer Applications (0975 – 8887) Volume 39– No.17, February 2012.

2. Shweta Bhatele and Lalita Bargadiya "Evaluation of Communication Overhead and Energy Consumption in Wireless Sensor Network using Different Clustering Techniques" IJSHRE, Vol. 2(1), 2014.

3. T. Mohana Priya, Dr. A.Saradh "An Improved K-means Cluster algorithm using Map Reduce Techniques to mining of inter and intra cluster datain Big Data analytics" International Journal of Pure and Applied Mathematics, Volume 119 No. 7 2018, pp. 679-690.

4. Nidhi Singh, Divakar Singh "Performance Evaluation of K-Means and Heirarichal Clustering in Terms of Accuracy and Running Time" International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012, pp. 4119-4121

5. Dibya Jyoti Bora, Dr. Anil Kumar Gupta "Effect of Different Distance Measures on the Performance of K-Means Algorithm: An Experimental Study in Matlab" Dibya Jyoti Bora et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, pp. 2501-2506.

6. Shaeela Ayesha, Tasleem Mustafa, Ahsan Raza Sattar & M.Inayat Khan, "Data Mining Model for Higher Education System ",European Journal of Scientific Research, ISSN 1450-216X Vol.43 No.1 ,2010, pp.27.

7. Amrinder Kaur, "Simulation of Low Energy Adaptive Clustering Hierarchy Protocol for Wireless Sensor Network", Volume 3, Issue 7, July 2013.

8. Amitabh Basu , Jie Gao, Joseph S. B. Mitchell, Girishkumar Sabhnani , "Distributed localization using noisy distance and angle information" 7th AC Minternational symposium on Mobile ad hoc networking and computing, 2006,pp. 262 – 273.

9. Rui Zhang, Myung J. Lee, Seong-Soon Joo, "Mobile sink update local information through these Aps", 2008.

10. Shiv Prasad Kori, "Performance Comparison in Terms of Communication Overhead for Wireless Sensor Network Based on Clustering Technique", Volume 4, Issue 3, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209.

11. Labisha R. V, Baburaj E, "Energy Efficient Clustering Algorithms in Wireless Sensor Networks-An Analytical View", Volume9, Number3, May 2014.

12. E. Ekici , S. Vural, J. McNair, D. Al-Abri "Secure probabilistic location verification in randomly deployed wireless sensor networks " Elsevier Science Publishers B. V, 2006 ,pp.,195-209