Evaluate the performance of K-Means and the fuzzy C-Means algorithms to formation balanced clusters in wireless sensor networks

Ali Abdul-hussian Hassan[1], Wahidah Md Shah[2], Mohd Fairuz Iskandar Othman[3], Hayder Abdul Hussien Hassan[4]
[1,2,3]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia
[1]College of Education for Pure Sciences, University of Kerbala, Iraq
[4]College of Management and Economics, University of Kerbala, Iraq

Article Info

ABSTRACT

The clustering approach is considered as a vital method for wireless sensor networks (WSNs) by organizing the sensor nodes into specific clusters. Consequently, saving the energy and prolonging network lifetime which is totally dependent on the sensors battery, that is considered as a major challenge in the WSNs. Classification algorithms such as K-means (KM) and Fuzzy C-means (FCM), which are two of the most used algorithms in literature for this purpose in WSNs. However, according to the nature of random nodes deployment manner, on certain occasions, this situation forces these algorithms to produce unbalanced clusters, which adversely affects the lifetime of the network. Based for our knowledge, there is no study has analyzed the performance of these algorithms in terms clusters construction in WSNs. In this study, we investigate in KM and FCM performance and which of them has better ability to construct balanced clusters, in order to enable the researchers to choose the appropriate algorithm for the purpose of improving network lifespan. In this study, we utilize new parameters to evaluate the performance of clusters formation in multi-scenarios. Simulation result shows that our FCM is more superior than KM by producing balanced clusters with the random distribution manner for sensor nodes.

Corresponding Author:

Ali Abdul-hussian Hassan,
Faculty of Information and Communication Technology,
UniversitiTeknikal Malaysia Melaka,
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.
Email: altaeeali800@yahoo.com

1. INTRODUCTION

Wireless Sensor Networks (WSNs) is utilized in numerous applications since they are suitable for various environments. It can function independently in conditions of harsh or hazardous places, where these places impose great risks to human beings, and is not advisable for them to be present there. Nevertheless, the sensor's lifetime is only related to their batteries, which are impossible to be replaced or recharged [1–3]. Consequently, with a view of prolonging the network lifetime, WSN used clustering approach for the clustering of the nodes, where the segregation of the sensor nodes into small clusters are executed based on their Euclidean distance. Each cluster employs one node to be the cluster head (CH). The CH possesses numerous functions in addition to sensing the environment such as; data gathering from all cluster members, and its conveyance to the main node termed as Base Station (BS), the conveyance of other CHs data to the next hop, and the fusion of the cluster data. Clustering approach is the most popular energy efficient technique which provides various advantages such as prolonging the network lifetime, scalability and

enabling less delay, where it is considered as an advantage for both the lifespan and the scalability of a network [4, 5]. In general, clustering algorithms are signified as the compilation of unsupervised classification methods that assigns objects into groups, or the partitioning of datasets into subsets known as clusters. Through the utilization of suitable clustering algorithm, the formation of clusters with objects that have the same features into the same cluster as opposed to objects in differing clusters is enabled. This means, clustering entails the allocation of objects possessing certain similarities into the same cluster according to their characteristics [6, 7]. One of the most important challenges faced by the clustering approach in WSN is how to improve the cluster structure and construct a balanced size of clusters [8]. Cluster size in our study refers to the quantity of member nodes in individual cluster. For this objective, several approaches were used based on KM and FCM algorithms for better clusters formation. Due to the nature of the random distribution of nodes in the monitoring area, at times these algorithms construct imbalanced clusters size [8], In this situation, large and small size of clusters are produced, as shown in Figure 1. Consequently, when the clusters sizes are not similar, the situation will lead to an imbalanced in the energy consumption among the nodes, which will result in a reduction in the lifespan of the network. Although there are many studies that have investigated on which of these algorithms is more superior for clustering process in other fields [6, 7, 9–13]. However, based on our knowledge, there is none that investigate which algorithm has a relatively better performance in terms forming a balanced size of clusters with the random distribution manner for nodes in the monitoring area. This is what motivated us to do an analysis study to investigate which of these algorithms has a better performance to form balanced clusters in the case of randomly node distribution. This study can enable researchers to choose the appropriate algorithm in order to improve network lifetime, where choosing an effective clustering method is the first issue that is faced during the construction of the clusters in the WSNs [14].

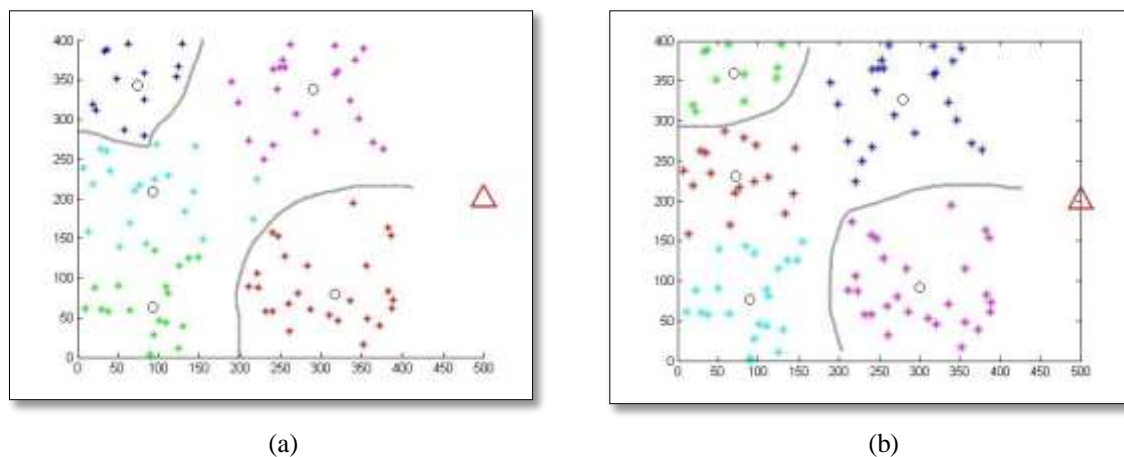(a)   (b)

Figure 1. Formation an imbalanced clusters size by (a) KM and (b) FCM

We have simulated several scenarios based on three measured parameters which are; Variation between clusters size, Standard deviation for Mean Square Error for intra-distances [15], and the ratio between minimum cluster size and maximum cluster size in the network. The remainder of the current study will be ensued by the ensuing sections; Section Two entails the related works. Additionally, in Section Three, we will explain the clustering algorithms. In Section Four, the scenarios and evaluation will be explained. Finally, Section 5 consists of the discussion and conclusion.

2. RELATED WORKS

Among the principle objectives in WSN is the effective clustering of the whole network, as it is able to decrease the energy being consumed [14] and also is able to offer balanced energy consumption. Hence, KM and FCM are the most utilized algorithms to realize this purpose. Initially, KM algorithm is used by researchers to construct balanced clusters sizes, where it increases the distances between the clusters along with the reduction distance inside the cluster, through the determination of the best cluster centroid. This is due to the fact that it is easily understood, and easily enforced, in addition to its uncomplicated features which resulted in this approach to be very popular. Another notable advantage of this method is that it does

not require prior calculation of the distances between the two entities that would have otherwise effect an extreme compromise in terms of operation duration and memory space. So, there are a few techniques which are based on the K-means algorithm used in clustering [16]. Ray and De [8] recommended an Energy-Efficient Clustering Protocol based on K-means midpoint algorithm (EECPK-means) for WSN. The midpoint algorithm is utilized to enhance the initial selection of the cluster centroid process. Moreover, the CH selection procedure will be optimized through the dependency on the residual energy as one of the parameters in the selection method as well as the Euclidean distance that is already utilized in KM. *As shown previously* by Mechta et al [17], they recommended Leach-CKM as an energy-effective method for the optimal energy usage. In their study, the researchers presented a pair of varying differing algorithms, which are the K-means algorithm for the purpose of clustering, and the routing protocol MTE for the transmission of data within the networks.

Through Bholowalia and Kumar [18], an additional version of Leach protocol was recommended for the balancing of energy usage by utilizing K-means and Elbow algorithm. *The underlying concept of the technique for the* acquisition is executed entailing the utilization of the k-means algorithm, and in ascertaining the optimum value "k," the Elbow technique was utilized. Consequently, a new cluster system for WSN was offered, with a dynamic system that automatically choose the amount of clusters. From Sheta et. al. [19], in their work, they recommended a hybrid protocol to extend the network lifespan. It utilized the k-means algorithms in clustering the nodes into groups. In addition, it incorporated the particle swarm optimization technique, and utilized the genetic algorithms for CH candidate's selection. Furthermore Rezaei et. al. [20], proposed a Multi-Hop Routing Energy Efficient Scheme (MRRCE). In their work, they improved the KM algorithm through the utilization of the Steiner Points (SPs) concept to enhance the selection of initial cluster centroid. Where they considered the SPs as an alternative way for a random selection of initial cluster centroid step in KM algorithm. The principle function of SPs is the fusing of each node with its neighbors, and the creation of a grid among nodes, where the KM algorithm will then select the initial cluster centroid based on this means rather than by random selection. Consequently, the CHs will determine the best location for the initial CH. In addition to that, Elkamel and Cherif [16] proposed a new technique in surmounting the energy usage issues. Their research aims were the incorporation of an enhanced algorithm, that entails the K-means in the generation of balanced energy in the clusters, in the addition of the utilization of the Gaussian elimination algorithm whilst the assignment of the CHs were being executed, which will ensure the dissemination of energy consumption. In order to resolve the issue of ascertaining the optimum quantity of groups, the researchers utilized the Davies Bouldin index to prolong the lifespan of the network. A *certain study by* Razzaq and Shin [21] recommended a system that takes into account K-means clustering during the clustering formation stage, and computes the weight function for the procedure of choosing the CH. Furthermore, it takes into account an optimal fixed packet size in relations to radio parameters and the state of the channel of of the transceiver. During the stage of the transference of the data, it applies a multi-objective weight function as a connection cost by utilizing the conventional Dijkstra algorithm.

Likewise, the FCM algorithm is widely used by researchers to construct balanced clusters size. Since it has the ability to determine the cluster's centroid such as KM, that helps in the optimization of the clusters according to the minimization of the space between the sensor node and the cluster centroid [22]. Alia [23] a Decentralized Fuzzy Clustering Protocol, named DCFP was suggested. The construction procedure of the framework for a particular WSNs is conducted one off at the starting of the protocol at a base station, that persists in its unaltered state transcending the entirety of the lifespan of the the network. At the beginning of the formation stage, a Fuzzy C-Means Algorithm is modified to assign the sensor nodes to their optimum suitable clusters. During the CH-Election stage, the assignment of new CHs is executed locally within individual cluster, in which instance, a new multi-criteria objective function is recommended for the enhancement of the quality of assigned cluster heads. Furthermore Bouyer et al [24] suggested a new method for minimizing energy consumption within the wireless sensor networks with hybrid LEACH protocol and Fuzzy C-Means Algorithm. The Fuzzy C-means (FCM) algorithm is utilized in the optimization of the number of the CHs and ascertaining their location and their allocation. The utilization of FCM in WSNs assists in changing the LEACH protocol parameters during the implementation. Through Hadjila et al [22] suggested a duo of algorithms utilizing a method which integrates the fFuzzy C-Means Algorithm and the ant colony optimization in the construction of the clusters, and the management of the data transference within the network. Firstly, Fuzzy C-Means clustering Algorithm is utilized in the formation of a predetermined amount of clusters. Secondly, the Ant Colony Optimization (ACO) algorithm was applied in the formation of a local minimal chain in individual clusters. In another research done by Kaushik [25] a hybrid approach based on Fuzzy C-Means clustering and neural network was suggested. The benefits of both methods, which are the Fuzzy C-Means clustering and neural network used to enable an energy effective network that prolonged the network lifespan had been utilized by the researcher. The formation of

the cluster is conducted through the utilization of FCM to construct evenly sized clusters within the network. Furthermore, the determination of CH selection is executed through the neural network, by taking into consideration the factors such as the proximity from the base station and the node energy. In their work, Shokrollahi et al [26] introduced an energy-efficient clustering algorithm founded on the Fuzzy C-Means Algorithm and genetic fuzzy system (ECAFG). Through the utilization of the FCM algorithm, the formation of clusters are conducted, followed by the selection of the CHs through utilization of a genetic fuzzy system (GFS). The formed clusters will continue to be unchanged, however the Cluster Heads are chosen at the starting of every turn. The FCM algorithm constructs balanced static clusters to decrease the data expenses, and disseminate the used energy amongst the clusters. Jain and Goel [27] proposed an Energy Efficient Routing Algorithm using Fuzzy C-Means (EEA-FCM). In that particular work, the Fuzzy C-Means clustering was utilized to form an optimal amount of static clusters. The notion of coherence was utilized to remove surplus and unneeded data generation, and unnecessary transmission which averts unwarranted energy wastage. The utilization of the Intra-cluster and inter-cluster gateways are to avert the nodes from transferring data over an extensive length.

Many researchers examined the performance of KM and FCM by functioned a relative comparison between these algorithms based on execution time and accuracy of these clustering algorithms in other fields [3, 4, 10, 12, 13, 28, 29]. Based on [11] KM is superior in terms the total sum of distances for nodes and time of execution in comparison to FCM when applied these algorithms on three intrusion datasets which are: KDDCup99, NSLKDD, and GureKDD. But in segmentation, the image of Brain Tumourin the [30], the error percentage value for image segmentation is lowest with FCM clustering and it outperforms KM. However, none of the research studied demonstrate which of these algorithms has more ability to form a balanced size of clusters under the random distribution manner of nodes in WSNs. This motivated us to make this study. In the next section, the details with pros and cons for each algorithm will be presented.

3.   CLUSTERING ALGORITHMS
3.1.   K-means algorithm (KM)

It is one of the unsupervised clustering methods, which efficiently utilized to form spherical shapes clusters. Stuart Lloyd was firstly researcher suggested this algorithm [31]. It divided points of data into a specific number of clusters [32]. It mostly increases the distances between the clusters along with reducing distance inside the cluster. The goal of this algorithm is sought to find the best cluster centroid when diminishing the objective function based on a Squared-Error-Function (SEF). Using the K-means algorithm, the clusters have a better formation where the average distance of each node of the cluster is minimized. It is more efficient to balance the load of the network to distribute the nodes between clusters [16]. This algorithm is very beneficial to construct the clusters for various applications of WSN [8]. The objective function of KM is defined as:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} d(x_i, x_c)^2 \quad i= 1, 2, ..., n \quad j=1, 2, ..., k \tag{1}$$

Where $d(x_i, x_c)^2$ represent the Euclidean distance that used to determine the distance between node x_{ij} with its cluster centroid c_j, i refers to numbea r of nodes, and j refers to cluster number. The processes of this algorithm are including the following phases [33]:

Phase 1: Locate the k centroids points in the space which is representing by the data set, where K is a predefined number.
Phase 2: Allocate every point of data to the specific cluster, which has the nearest centroid distance.
Phase 3: Once all point of data has been clustered, re-determine the locations of the k centroids.
Phase 4: reiterate the Phase 2 and Phase 3 till no shown change in the location of centroids.

3.2.   Fuzzy C-Means (FCM)

It is considered as one of the most efficient protocols [34], in many real situations, the fuzzy clustering methods, dealing with uncertainty, fuzziness, and vagueness, fuzzy clustering is considering as an effective clustering method. Among the fuzzy clustering methods, the fuzzy C-means (FCMs) algorithm has been most widely used in the clustering processes [35]. The goal of FCM is to minimize the sum of distances between the instances and the cluster centers [36]. In WSNs, the aim is to cluster N sensor nodes into k distinguished clusters. The objective function of FCM for clustering in WSNs can be formulated as follows:

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} \mu^m \, d(x_i, x_c)^2 \,, \quad i= 1, 2, ..., n \quad j=1, 2, ..., k \tag{2}$$

$$\mu_{ij} = \frac{1}{\sum_{j=1}^{k} (\frac{d(x_i,c_j)}{d(x_i,c_k)})^{\frac{2}{m-1}}} \qquad (3)$$

$$\mu_{ij} \in [0, 1]$$

$$C_j = \frac{\sum_1^n (\mu_{ij})^m x_i}{\sum_1^n (\mu_{ij})^m} \qquad (4)$$

Where μ is the membership of node i to cluster j, m is the value of fuzzifier is usually chosen as 2 in the most of applications [7]. And C_j refers to cluster centroid. This function differs from KM with the use of weighted squared errors instead of using squared errors only. These clustering algorithms still have some disadvantages that hinder its function, Table 1 shown the disadvantages for both KM and FCM.

Table 1. Disadvantages of KM and FCM

	KM	FCM
1	The initial centroids are selected by the random way for the input data set.	Same like KM
2	sensitivity to outliers points	Same like KM
3	The number of clusters K is given as manually.	The number of clusters K and the fuzzy weighted index (m) is given as manually
4	No guarantee for K-means will converge into an optimal or better solution.	It relapses into the local extreme point or saddle point easily and the optimal solution cannot be obtained.
5	In the clusters formation, size of clusters is not considering	Same like KM

4. SCENARIOS AND EVALUATION

We examined totally six scenarios for comparing KM and FCM using Matlab and excel. Each scenario has different 50 observations i.e. different 50 distribution pattern (uniform random distribution) for nodes in the monitoring area. Although, the study [8] relied on 7 observations in their evaluation in order to decide whether the resulting clusters are balanced or not. However, to increase the accuracy of our evaluation result, our study relied on different 50 observations.

To evaluate the performance for these algorithms in the formation of a balanced size of clusters, we utilize three norms which are:

1. Standard deviation of Mean Square Error STD (MSE) for intra-distances: Which measures what the difference in the homogeneity for the average of intra-distance for each cluster. This norm shows how the average intra-distances of nodes to the cluster's centroid are different from cluster to others. Where the smaller the factor, the better. That mean, there is a uniformity of the intra-distances for clusters.

$$STD\ (MSE) = \sqrt[2]{\frac{\sum |MSE(j)-\mu|^2}{k}} \qquad (5)$$
$$J=1, 2, ..., k$$

Where $STD\ (MSE)$ is mean the standard deviation of Mean Square Error, k is the number of clusters, and μ is the Average of Mean Square Error for distances.

$$MSE\ (j) = (1/n) * \sum_{i=1}^{n} D(x_i, x_c)^2 , \qquad (6)$$
$$i=1, 2,..., n \qquad c=1, 2,..., k$$

The acronym MSE refers to the average of square intra-distances of nodes to the cluster's centroid, n is number of nodes in each cluster, and $D(x_i, x_c)^2$ square intra distances for node (x_i) to its cluster centroid (x_c) in the cluster (c).

$$\mu = \frac{\sum_{j=1}^{k} MSE}{k} \qquad (7)$$

2. Variation for clusters size (V): Which measures the dissimilarity of the density of the nodes in the clusters (number of member nodes in each cluster). Where the smaller the factor, the better. That mean, there is a balance in clusters size.

$$V = \frac{\sum |S_j - \mu|^2}{k} \qquad (8)$$

$$\mu = \frac{\sum_{j=1}^{k} S_j}{k} \qquad (9)$$

Where S_j refers to cluster size (j) and μ refer to the mean of clusters size.

3. Clusters Size Range (CSR): Which measures the ratio of minimum cluster size to maximum cluster size. So, the range of clusters size be limited within this range (CSR to 1), the narrower the range the better (close to 1). That means no big difference in size between minimum cluster size to maximum cluster size.

$$CSR = \min\left(\frac{CS_j}{max.CS}\right) \qquad (10)$$

Where CS_j refers to clusters size and max.CS refers to the maximum cluster size in the network.

There are many parameters that utilized in the literature in WSN such as (residual energy, distance to BS, velocity, and etc.). However from perspective of formation a balanced size of clusters in network, only these novel parameters based on (intra-distance, nodes density of cluster, and clusters size) shown whether it is a balanced cluster's size or not. These new parameters are considering as our main contribution to this study.

In this study, we utilize these parameters together, where it is not sufficient to consider this network has a balanced size of clusters more than others networks depending only on the density of the distribution of the nodes within the clusters, regardless of the homogeneity to the average intra-distances in the clusters, vice versa. In addition, is very essential to determine the range of clusters size, where the volumetric width shows the difference in size between the largest and smallest cluster in the network. The movement in cluster sizes is from CSR value to 1, CSR value the narrower (close to 1) the better. So, these three parameters used to evaluate the performance of these algorithms, in terms which of these algorithms have the ability to produce more balanced clusters compared to another with the random distribution manner for nodes in the monitoring area for WSN.

In this study, we depended on the most frequent scenarios in the literature. We applied KM and FCM for different 50 uniformly random distributions to divide nodes into five clusters for each scenario when the BS location (x, y) is located outside network. Also, based on literature Squared Euclidean distance norm was used as the distance measure in both KM and FCM algorithms, the remaining details of these scenarios illustrated in Table 2.

Table 2. Details of scenarios

Number of scenarios	Number of nodes	Monitoring area (m)
1st scenario	100	200*200
2nd scenario	100	400*400
3rd scenario	100	500*500
4th scenario	200	200*200
5th scenario	200	400*400
6th scenario	200	500*500

5. EVALUATION AND RESULTS

Firstly, we evaluated the performance of KM and FCM to construct balanced clusters size based on every parameter separately in each scenario. According STD (MSE) parameter, KM outperformed on FCM 19, 20, 10, 17, 22, and 19 observations (distributions) from 50 distributions in first scenario to sixth scenario, respectively. Nonetheless FCM superior KM 31, 30, 40, 33, 28, and 31 observations from 50 observations for each scenario from the first scenario to the sixth scenario, respectively. Which mean 62%, 60%, 80%, 66%, 56%, and 62% for each scenario from first scenario to sixth scenario, respectively. According to the CSR parameter, KM outperformed on FCM 17, 16, 16, 9, 10, and 8 observations from 50 distributions for each scenario from the first scenario to the sixth scenario, respectively. But FCM superior KM 32, 33, 30, 39, 37, and 42 observations from 50 observations for each scenario from the first scenario to the sixth scenario, respectively. Which mean 64%, 66%, 60%, 78%, 74%, and 84% for each scenario from first scenario to sixth scenario, respectively. In addition, 1, 1, 4, 2, and 3 similar value of CSR for both KM and FCM in first, second, third, fourth, and fifth, respectively.

According to the variation parameter, KM outperformed on FCM 17, 17, 13, 10, 11, and 10 observations from 50 distributions in the first scenario to the sixth scenario, respectively. However, FCM superior KM 33, 33, 35, 39, 38, and 40 observations from 50 observations for each scenario from the first scenario to the sixth scenario, respectively. Which mean 66%, 66%, 70%, 78%, 76%, and 80% for each scenario from first scenario to sixth scenario, respectively. In addition, 2, 1, and 1 similar value of variation for both KM and FCM in the third, fourth, and fifth scenario, respectively. Figure 2 shows the result for each parameter in different six scenarios.
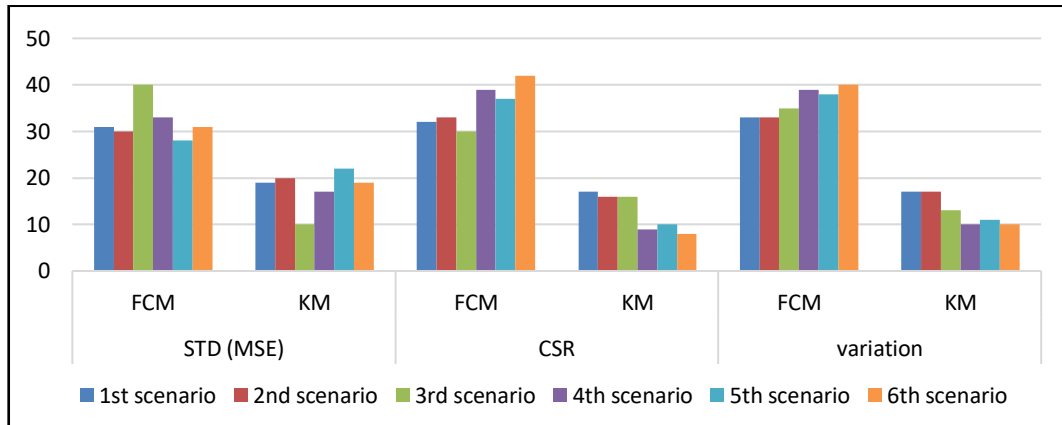

Figure 2. The performance-based three parameters in different six scenarios

Secondly, we evaluate KM and FCM based on scenarios in terms which of these algorithms has better performance based on these three together parameters (3 parameters) in the same observation or distribution for all scenarios. The first scenario, KM is succeeded by 8 observations out of different 50 observations, while FCM is better than KM in 22 observations, where it achieved all together parameters in 22 observations out of different 50 observations. In the second scenario, although KM is succeeded by 8 observations out of different 50 observations, FCM is better than KM in 21 observations, where it achieved all together parameters in 21 observations out of different 50 observations. In the third scenario, KM is succeeded by 3 observations out of different 50 observations, whereas FCM is better than KM in 26 observations, where it achieved all together parameters in 26 observations out of different 50 observations. In the fourth scenario, KM is succeeded by 3 observations out of different 50 observations, while FCM is better than KM in 26 observations, where it achieved all together parameters in 26 observations out of different 50 observations. In the fifth scenario, KM is succeeded by 5 observations out of different 50 observations, however, FCM is better than KM in 21 observations, where it achieved all together parameters in 21 observations out of different 50 observations. In the last scenario, KM is succeeded only in 1 observation out of different 50 observations, while FCM is better than KM in 24 observations, where it achieved all together parameters in 24 observations out of different 50 observations. Consequently, based on the three together parameters, FCM is better than the KM in terms construct balanced clusters size as shown in Figure 3.
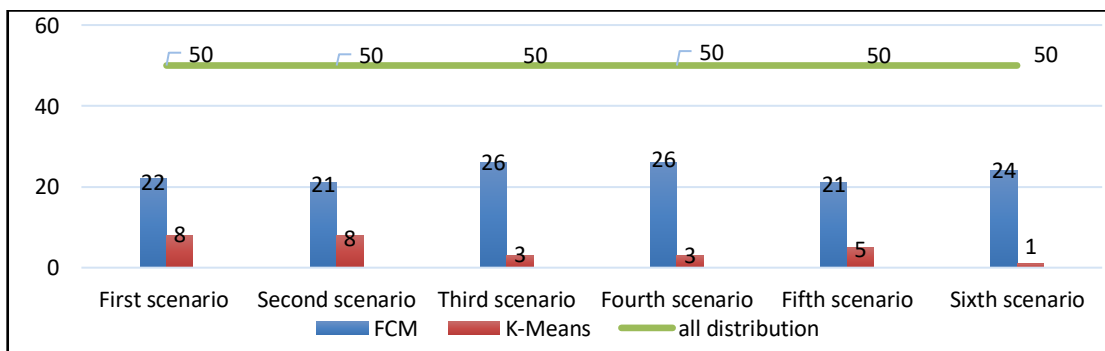

Figure 3. Formation of balanced clusters based on three together parameters

6.   DISCUSSION AND CONCLUSION

In this study we evaluated the performance of KM and FCM in terms construct balanced clusters size for random nodes deployment in the monitoring area in WSN. The evaluation was based on six scenarios, each scenario has different 50 observation of nodes and the nodes in each observation have divided into five clusters. Also, we utilized a Squared Euclidean distance norm as the distance measure in both KM and FCM algorithms. Three new parameters associated with the size of the cluster has been used in this evaluation, which is Variation between clusters size, Standard deviation for Mean Square Error for intra-distances, and the ratio between minimum cluster size and maximum cluster size in the network. Based on the result, FCM has better performance than KM to formation a balanced cluster's size based on these parameters with the random distribution manner for sensor nodes in the monitoring area. Also, when the number of nodes distributed is increasing along with the increase in the monitoring area, the performance of FCM still relatively stable compared to KM, where the performance of KM decreased. Although FCM is superior to KM, but still suffer from the effect of the random nodes deployment condition, where sometimes form imbalanced clusters. This limitation requires to propose assist mechanism to overcome this problem, this will be addressed in futurework. At the conclusion and based on the result, FCM is a better choice to form a balanced cluster especially when the number of nodes distributed is high along with the big distance of the monitoring area in random nodes deployment in the monitoring area for WSNs.

REFERENCES

[1]   E. C. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, "Wireless sensor networks: a survey," *Comput. Networks*, vol. 38, no. 4, pp. 393–422, 2002.
[2]   A. A. Hassan, W. Shah, M. F. Iskandar, and M. N. Al-mhiqani, "Unequal Clustering Routing Algorithms in Wireless Sensor Networks : A Comparative Study," *Journal of Advanced Research in Dynamical and Control Systems,* vol. 10, pp. 2142–2156, 2018.
[3]   S. Al-Khammasi, D. Alhelal, and N. S. Ali, "Energy efficient cluster based routing protocol for dynamic and static nodes in wireless sensor network," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 16, no. 5, pp. 1974–1981, 2018.
[4]   P. Maratha and P. Kapil, "A Comparative Study on Prominent Strategies of Cluster Head Selection in Wireless Sensor Networks," *Integrated Intelligent Computing, Communication and Security*, Springer Singapore, 2019.
[5]   F. Liang, L. Zhang, and P. Sun, "Study on the Rough-set-based Clustering Algorithm for Sensor Networks," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 3, no. 2, pp. 77–90, 2014.
[6]   S. Al-Augby, S. Majewski, A. Majewska, and K. Nermend, "A Comparison Of K-Means And Fuzzy C-Means Clustering Methods For A Sample Of Gulf Cooperation Council Stock Markets," *Folia Oeconomica Stetin.*, vol. 14, no. 2, pp. 19–36, 2014.
[7]   Z. Cebeci and F. Yildiz, "Comparison of K-Means and Fuzzy C-Means Algorithms on Different Cluster Structures," *J. Agric. Informatics*, vol. 6, no. 3, pp. 13–23, 2015.
[8]   A. Ray and D. De, "Energy efficient clustering protocol based on K-means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network," *IET Wirel. Sens. Syst.*, vol. 6, no. 6, pp. 181–191, 2016.
[9]   S. W. Guangul, "The effects of segmentation techniques in digital image based identification of ethiopian paper currency," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 3, pp. 1106–1110, 2018.
[10] A. Kapoor, "A Comparative Study of K-Means , K-Means ++ and Fuzzy C- Means Clustering Algorithms," *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT)*, pp. 1–6, 2017.
[11] S. K. Sahu and S. K. Jena, "A study of K-Means and C-Means clustering algorithms for intrusion detection product development," *Int. J. Innov. Manag. Technol.*, vol. 5, no. 3, 2014.
[12] T. Singh and M. Mahajan, "Performance Comparison of Fuzzy C Means with Respect to Other Clustering Algorithm," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no. 5, pp. 89–93, 2014.
[13] A. Sheshasayee and P. Sharmila, "Comparative Study of Fuzzy C Means and K Means Algorithm for Requirements Clustering," Indian Journal of Science and Technology, vol. 7, no. 6, pp. 853–857, 2014.
[14] S. Dehghani and B. Barekatain, "An Enhanced Energy-Aware Cluster-Based Routing Algorithm in Wireless Sensor Networks," *Wirel. Pers. Commun.*, vol. 98, no. 1, pp. 1605–1635, 2018.
[15] H. Karim, S. R. Niakan, and R. Safdari, "Comparison of neural network training algorithms for classification of heart diseases," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 7, no. 4, pp. 185–189, 2018.
[16] R. ELkamel and A. Cherif, "Energy-efficient routing protocol to improve energy consumption in wireless sensors networks," *Int. J. Commun. Syst.*, vol. 30, no. 17, pp. e3360, 2017.

[17] D. Mechta, S. Harous, I. Alem, and D. Khebbab, "LEACH-CKM: Low Energy Adaptive Clustering Hierarchy protocol with K-means and MTE," in *2014 10th International Conference on Innovations in Information Technology (IIT)*, pp. 99–103, 2014.

[18] K. Bholowalia P, "EBK-Means:Aclustering technique based on elbow method and k-means in WSN," *Int J Comput Appl.*, vol. 105, no. 9, 2014.

[19] A. F. Sheta and B. Solaiman, "Evolving clustering algorithms for wireless sensor networks with various radiation patterns to reduce energy consumption," in *2015 Science and Information Conference (SAI)*, pp. 1037–1045, 2015.

[20] E. Rezaei, A. A. Baradaran, and A. Heydariyan, "Multi-hop Routing Algorithm Using Steiner Points for Reducing Energy Consumption in Wireless Sensor Networks," *Wirel. Pers. Commun.*, vol. 86, no. 3, pp. 1557–1570, 2016.

[21] M. Razzaq and S. Shin, "Energy Efficient Dijkstra-Based Weighted Sum Minimization Routing Protocol for WSN," *2018 Third International Conference on Fog and Mobile Edge Computing (FMEC)*, Jun 2018.

[22] M. Hadjila *et al.*, "A Hybrid Cluster and Chain-based Routing Protocol for Lifetime Improvement in WSN," *International Conference on Wired/Wireless Internet Communications*, 2015.

[23] O. M. D. Alia, "A decentralized fuzzy c-means-based energy-efficient routing protocol for wireless sensor networks," *Sci. World J.*, vol. 2014, 2014.

[24] A. Bouyer, "A New Approach for Decreasing Energy in Wireless Sensor Networks with Hybrid LEACH Protocol and Fuzzy C-Means Algorithm," *Int. J. Commun. Networks Distrib. Syst.*, Nov 2014.

[25] A. K. Kaushik, "A Hybrid Approach of Fuzzy C-means Clustering and Neural network to make Energy-Efficient heterogeneous Wireless Sensor," *International Journal of Electrical and Computer Engineering,* vol. 6, no. 2, pp. 674, Apr 2016.

[26] Ayub Shokrollahi and Babak Mazloom-Nezhad Maybodi, "An Energy-Efficient Clustering Algorithm Using Fuzzy C-Means and Genetic Fuzzy System for Wireless Sensor Network," *Journal of Circuits, Systems and Computers*, vol. 26, no. 1, pp. 1–22, 2017.

[27] A. Jain, "Energy Efficient Algorithm for Wireless Sensor Network using Fuzzy C-Means Clustering," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 9, no. 4, pp. 474–481, 2018.

[28] D. J. Bora, "A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 10, no. 2, pp. 108–113, 2014.

[29] Sanjay Kumar and Ghosh, Soumi, "Comparative analysis of k-means and fuzzy c-means algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 4, 2013.

[30] H. Hooda and O. P. Verma, "Brain Tumor Segmentation : A Performance Analysis using K-Means, Fuzzy C-Means and Region Growing Algorithm," *International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, May 2014.

[31] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[32] G. A. Wilkin and H. Xiuzhen, "K-means clustering algorithms: Implementation and comparison," in *Proceedings - 2nd International Multi-Symposiums on Computer and Computational Sciences, IMSCCS'07*, 2007, pp. 133–136.

[33] T. M. Kodinariya, "Review on determining number of Cluster in K-Means Clustering," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 1, no. 6, pp. 90–95, 2013.

[34] D. Thanh, L. Hoang, and V. Trong, "Novel fuzzy clustering scheme for 3D wireless sensor networks," *Appl. Soft Comput. J.*, vol. 54, pp. 141–149, 2017.

[35] C. Lu, S. Xiao, and X. Gu, "Improving fuzzy C-means clustering algorithm based on a density-induced distance measure," *The Journal of Engineering,* Jan 2016.

[36] V. Krishnaswamy and S. K. S. Manvi, "Clustering and data aggregation scheme in underwater wireless acoustic sensor network," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 17, no. 4, pp. 1604–1614, 2019.