Optimal Gateway Placement In Iot Sensor Network Using Enhanced Fuzzy C-Means With Elbow Method Of Optimal Cluster Determination

[1]Ifeagwu E.N.
Department of Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State.
ORCID:0009-0005-7448-0187
ifeagwuen@fuotuoke.edu.ng

[2] Geku Diton
Department of Electrical and Electronic Engineering
Federal University Otuoke, Bayelsa State.
Email: gekudd@fuotuoke.edu.ng

[3] Ezema D.C.
Department of Electrical and Electronic Engineering
[3]State University of Medical and Applied Sciences(SUMAS),Igbo-Eno,Enugu State, Nigeria
donatus.ezema@sumas.edu.ng

***Abstract*— In this paper, the optimal gateway placement in IoT sensor network using Enhanced Fuzzy C-Means (E-FCM) with Elbow method of optimal cluster determination is presented. The essence of this paper is to employ Elbow method to conduct optimal number of cluster selection for an IoT sensor network and then use Enhanced Fuzzy C-means (E-FCM) algorithm to determine the optimal placement of the gateways within the clusters and then cluster the sensor nodes among the selected gateways. In the paper, 2000 sensor nodes spread over a rectangular area of 1000 m by 1000 m are considered. The results from the Elbow method show that the optimal number of clusters required for the sensor network is 5. Also, the results from the E-FCM show that cluster 1 has the highest number of nodes (463 nodes) which is about 23.2% of the total number of nodes in the network. On the other hand, cluster 0 has the lowest number of nodes (353 nodes) which is about 17.7% of the total number of nodes in the network. The average number of nodes per cluster is 20. Again, cluster 1 has the highest mean distance of 194 m per node which is about 110 % of the overall mean distance of 176 m for all the nodes in the network. On the other hand, cluster 0 has the lowest mean distance of 161 m per node which is about 92 % of the overall mean distance of 176 m. The highest distance to getaway occurred in cluster 1 with a distance of 334 m.**

Keywords*— *Gateway Placement, Internet of Things, Sensor Network, Enhanced Fuzzy C-Means, Elbow Method

1. Introduction

Nowadays, IoT sensor networks are deployed in diver smart systems applications. From the smart home to smart agriculture, smart city, smart transportation, smart grid and many other smart systems (Rajkumar and Santhosh Kumar, 2024; Chataut, Phoummalayvane and Akl, 2023; Bellini, Nesi and Pantaleo, 2022).In all these applications, the sensors play very vital role of capturing the environmental parameters or the system parameters upon which the smart systems depend for the decision making(Shen, Liu, Tian and Na, 2022; Sarker,2021). In each application, wireless communication of the sensor nodes and the base station or gateway is inevitable (Zhang, Zhang, 2022; Deniz, Bagci, Korpeoglu and Yazıcı,2021).As such, effective and efficient getaway-sensor node communication is required for the IoT sensor networks.

Researchers have shown that the energy expanded in any wireless communication is directly proportional with the communication distance (Saunders and Aragón-Zavala, 2024; Pizzo, Sanguinetti and Marzetta, 2023; Islam, Ahmad, Habibi and Waqar, 2022). Also, in many cases the sensor nodes are battery powered with finite life span (Bathre and Das, 2022; Abner, Wong and Cheng, 2021). As such energy efficient system is required in the IoT sensor network (Kaur, Chanak and Bhattacharya, 2021; Shaker Reddy and Sucharitha, 2022). One way to minimize the energy consumption in sensor networks is by the use of clustering with optimal getaway placement (Shanmugam and Kaliaperumal, 2021; El Ouadi, 2022).

In practice, the optimum number of clusters for a given network is required before the getaway placement is conducted. Accordingly, in this study, the Elbow method is used for the determination of the optimum number of cluster while the enhanced fuzzy C-means is used for the optimal gateway location determination for each of the clusters in the network (Hamka and Ramdhoni, 202). With this approach, network designers can achieve optimal energy consumption and hence enhanced network battery life span.

2. Methodology

The essence of this paper is to employ Elbow method to conduct optimal number of cluster selection for an IoT sensor network and then use Enhanced Fuzzy C-means (E-FCM) algorithm to determine the optimal placement of the gateways within the clusters and then cluster the sensor nodes among the selected gateways. The research procedure used is presented in Figure 1.

The data set used consists of N number of sensor nodes randomly distributed within the network area of coverage specified in terms of the dimensions of a rectangle with L as the length and W as the width, both specified in kilometers. The preprocessing of the dataset assessed the dataset for missing data, outliers. Also the sensor node coordinates were normalized with the minmax method.

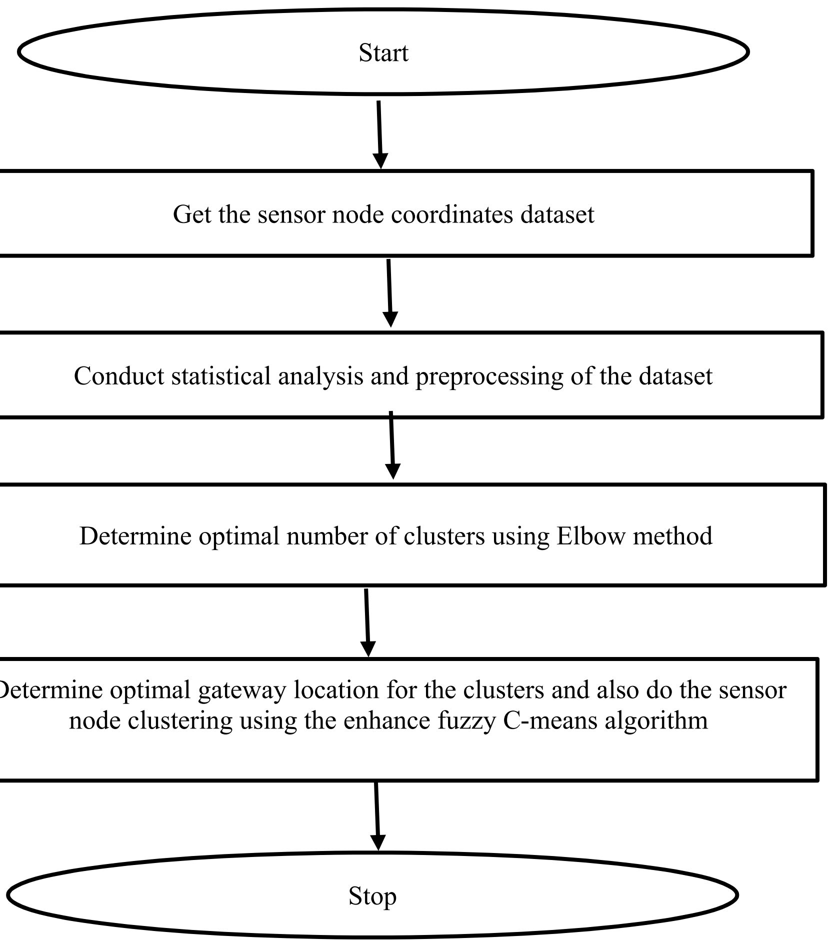


Figure 1: The research procedure

2.1 The optimal number of clusters for the sensor network using the Elbow Method

The optimal number of clusters for the sensor network was determined using the Elbow method. The Elbow method utilized Within-Cluster Sum of Squares (WCSS) score plot in Equation 1 to determine the optimum value of k which is the number of clusters.

$$WCSS = \sum_{i=1}^{i=k} \left(\sum_{x \in C_i}^{N_{C_i}} (\|x - \mu_i\|^2) \right) \qquad (1)$$

where K is used to represent the number of clusters, C_i is used to represent the ith clusters, x used to represent the data point (or sensor node x location coordinate) in clustersC_i, μ_i is used to represent the centroid of cluster C_i and N_{C_i} is used to represent the number of data points (or sensor nodes) in cluster C_i. The flow chart for the Elbow method is shown in Figure 2.

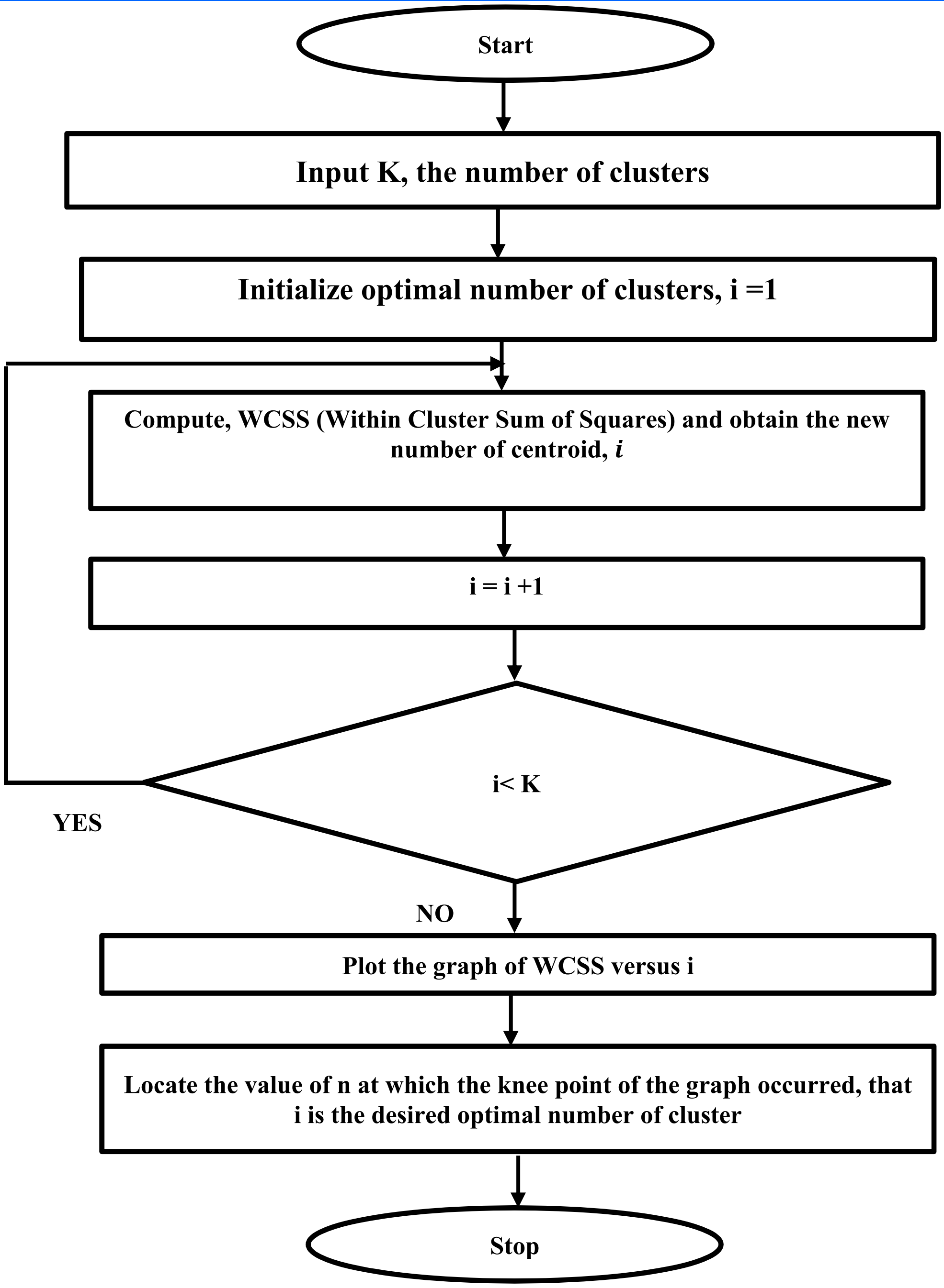


Figure 2: The flow chart for the optimum number of cluster determination using the Elbow method

2.2 Determination of Optimal Gateway Location using E-FCM Algorithm

Fuzzy C-Means seems to be a good choice for handling ambiguity. However, it is very slow when the dataset is very large. This slowness is due to the number of iterations it must complete to obtain the optimal solution. The number of iterations cannot be predetermined. This research proposes an enhanced method based on Fuzzy C-Means to expedite the iteration process. The new method uses the grid based approach to determine the gateway location.

Consider a cluster C_i where a grid is defined on the area covering the sensor nodes. This research adopts the local defect optimization approach to determine the region where the cluster head optimally lies. The approach precisely incorporates the solutions within a local region of the grid to the extended or global region of the grid. First consider a boundary value problem stated in Equation2 (Bockelmann *et al.*, 2019).

$$\xi\tau = f \in \omega \qquad (2)$$

where, ξ is a random concise operator, τ denotes a function on the boundary, and f denotes a function on the grid composite grid g. The grid composite $g^{h,H}$ constitutes the global region and the local region.

In this case, a local region g^{h} contains a single cell while a global region g^{H} is a conglomerate of single neighboring cells. Hence the global region can be defined on the (x, y) plane as expressed in Equation 3;

$$g^{H} = \left\{ (x_i, y_j) \mid x_i = iH,\ y_j = jH,\ 0 > i > \frac{1}{H},\ 0 < j < \frac{1}{H} \right\} \quad (3)$$

where g^{H} denotes the local region, x_i denotes the horizontal plane, y_j denotes the vertical plane, H denotes the global region identity.Similarly, it is expressed in Equation 4.

$$g^{h} = \left\{ (x_i, y_j) \mid x_i = ih,\ y_j = jh,\ 0 > i > \frac{1}{h},\ 0 < j < \frac{1}{h} \right\} \quad (4)$$

where h denotes the local region identity. The region of congestion of the specified cluster lies within g^{h}, and $g^{h,H}$ is defined as shown in Equation 5 as follows:

$$g^{h,H} = g^{h} \cup g^{H} \quad (5)$$

Hence, the cell integration factor η can be computed as shown in Equation 6.

$$\eta = \frac{H}{h} \quad (6)$$

Literarily, the integration of cell to obtain the global congested region is scaled by η. The prerequisite to the implementation of local defect optimization is the discretization of the boundary value problem on the extended grid which can be given as in (Bockelmann*et al.*, 2019)

$$\xi^{H} \tau_0^{H} = f^{H} \quad (7)$$

where, ξ^{H} denotes the random concise operator on the global region. The τ_0^{H} component which denotes the grid function is the approximate solution. This solution is used to compute the integral grid approximation as expressed in Equation8;

$$\tau_0^{H,h}(x) = \begin{cases} \tau_0^{h}(x) & x \in g^{h} \\ \tau_0^{H}(x) & x \in \frac{g^{H,h}}{g^{h}} \end{cases} \quad (8)$$

The global grid function as an extract of $\tau_0^{H,h}(x)$ can be expressed as in Equation9.

$$G_0^{H}(x) = \begin{cases} \tau_0^{h}(x) & x \in g^{h} \\ \tau_0^{H}(x) & x \in \frac{g^{H}}{g^{h}} \end{cases} \quad (9)$$

Hence, by adopting the grid based approach to determine the gateway location the E-FCM method minimizes the number of iterations required especially when the number of data points are many. This approach therefore reduces the time it takes to execute the gateway placement for large sensor networks.

3. Results and Discussion

In the paper, 2000 sensor nodes spread over a rectangular area of 1000 m by 1000 m are considered. The scatter plot of the sensor nodes randomly distributed in the network is presented in Figure 3 whereas the descriptive statistics of the sensor nodes X and Y coordinates is presented in Table 1. The descriptive statistics shows that there is no missing data and no outliers. Also, in Table 1, 2000 data points with maximum X and Y coordinates values of 1000 m are noted with the X coordinate having higher mean of 499.491 m while the Y coordinate has mean value of 485.315 m.

Again, the Within-Cluster Sum of Squares (WCSS) versus number of clusters plot for the Elbow method is presented in Figure 4. The results in Figure 4 show that the optimal number of clusters required for the sensor network is 5.

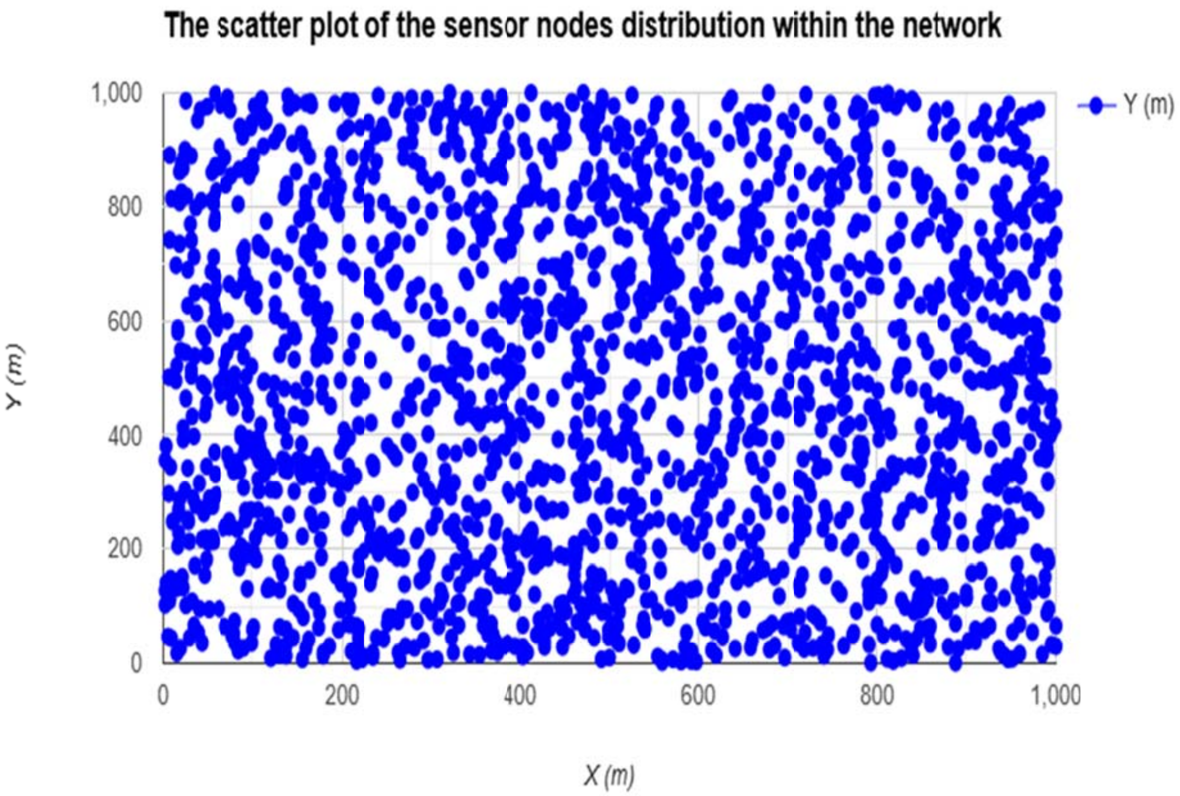


Figure 3: The scatter plot of the sensor nodes randomly distributed in the network

Table 1: The descriptive statistics of the sensor nodes X and Y coordinates

Groups	X (m)	Y (m)
Number of observations	2,000	2,000
Number of missing values	0	0
Minimum	2	1
Maximum	1,000	1,000
Range	998	999
Mean ($\bar{x}$)	499.491	485.315
Standard Deviation (S)	291.0598	292.4188
Q1	248	233.5
Median	492	479
Q3	749.5	737
Interquartile range	501.5	503.5
Outliers	None	None

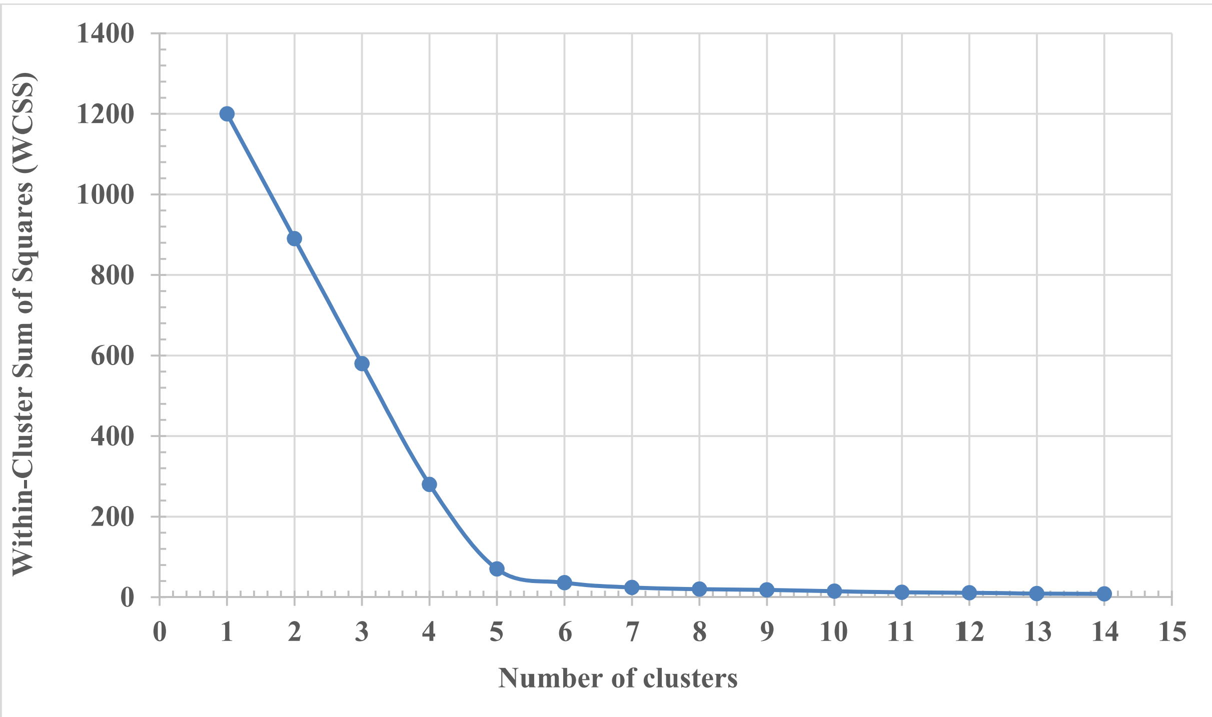
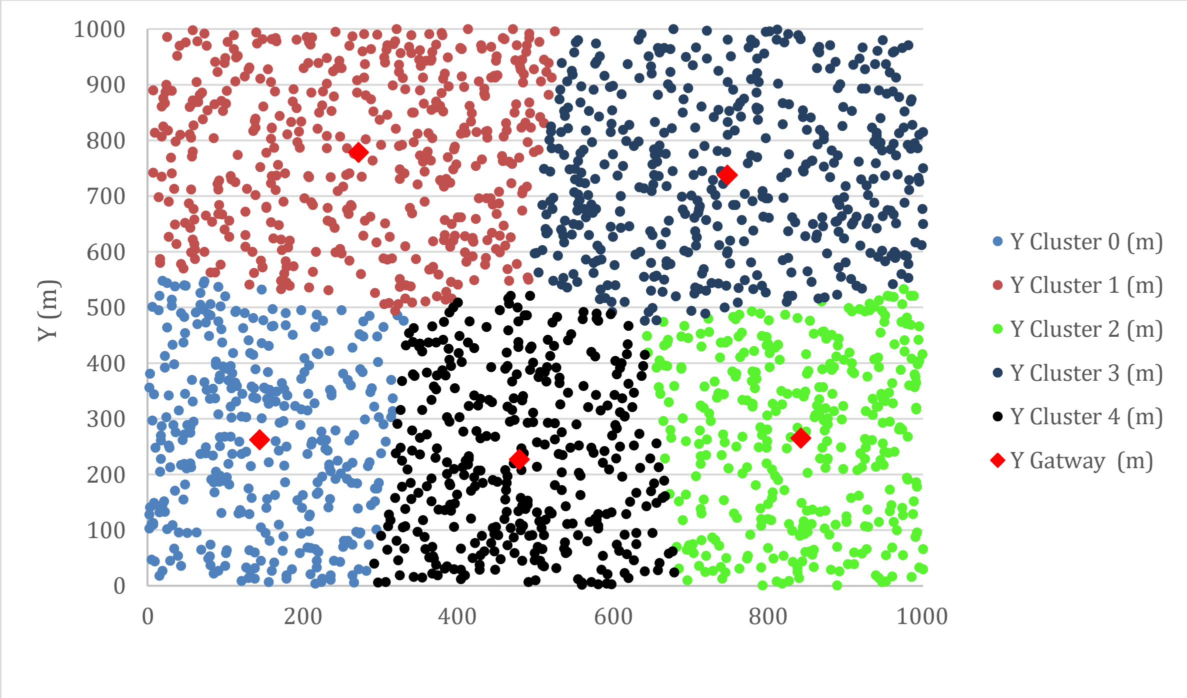


Figure 4: The Within-Cluster Sum of Squares (WCSS) versus number of clusters plot for the Elbow method



-Figure 5: Thegateway placement and clustering using the enhanced fuzz C-means algorithm

Table 2: The gateway X and Y coordinates determined using the fuzzy C-Means algorithm

Coordinated of the gateways for the 5 clusters		
	The gateway (centroid) X coordinate in m	The gateway (centroid) Y coordinate in m
Cluster 0	144.2181	262.4079
Cluster 1	271.9093	778.4903
Cluster 2	842.2238	265.232
Cluster 3	747.5	737.7022
Cluster 4	479.1271	227.0773

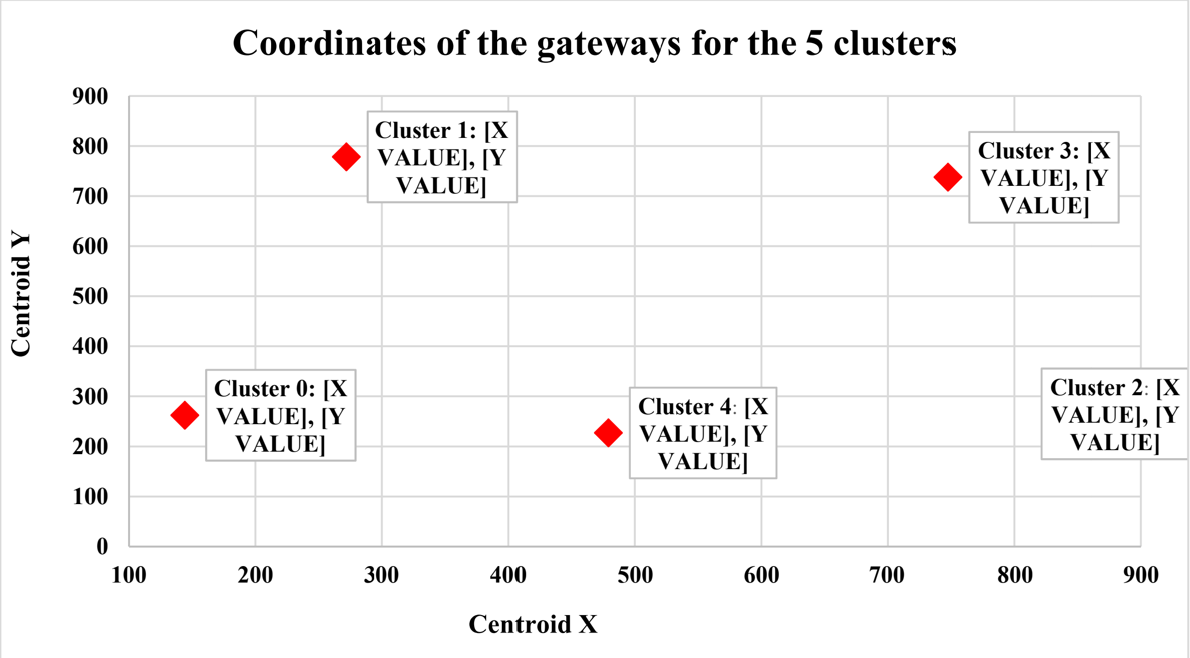


Figure 6: The scatter plot of the X and Y coordinates of thegateways

Table 3: The summary of the number of nodes and distance to the gateways for the sensor nodes in the network

	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Min	Max	Mean
No. of nodes	353	463	362	460	362	353	463	400
Percentage of total nodes	17.7	23.2	18.1	23.0	18.1	17.7	23.2	20.0
Minimum Distance from the Gateway (m)	4	6	13	9	10	4	13	9
Maximum Distance from the Gateway (m)	312	334	299	330	294	294	334	314
Average Distance from the Gateway (m)	162	194	168	191	165	162	194	176

The results of the gateway placement using the E-FCM algorithm is shown in Figure 5, Table 2 and Figure 6.The summary of the number of nodes and distance to the gateways for the sensor nodes in the network is shown in Table 3. The bar chart of the number of sensor nodes per cluster as determined using the enhanced fuzzy C-means is presented in Figure 7 while the bar chart of the mean distance of the sensor nodes from the gateway as determined using the enhanced fuzzy C-means is presented in Figure 8. According to the results, cluster 1 has the highest number of nodes (463 nodes) which is about 23.2% of the total number of nodes in the network. On the other hand, cluster 0 has the lowest number of nodes (353 nodes) which is about 17.7% of the total number of nodes in the network. The average number of nodes per cluster is 20.

In addition, the results in Table 3 and Figure 8 for the mean distance of the sensor nodes from the gateway showed that cluster 1 has the highest mean distance of 194 m per node which is about 110 % of the overall mean distance of 176 m for all the nodes in the network. On the other hand, cluster 0 has the lowest mean distance of 161 m per node which is about 92 % of the overall mean distance of 176 m. The highest distance to getaway occurred in cluster 1 with a distance of 334 m, as shown in Figure 8.

Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 10 Issue 2, February - 2025

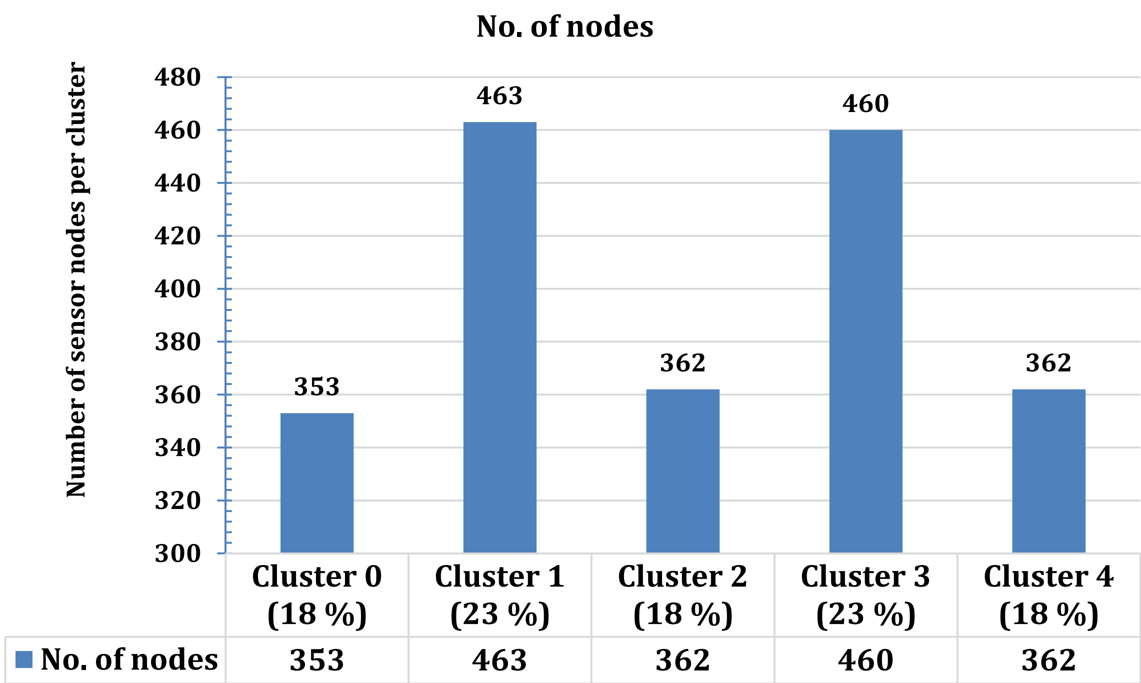


Figure 7: The bar chart of the number of sensor nodes per cluster as determined using the enhanced fuzzy C-means

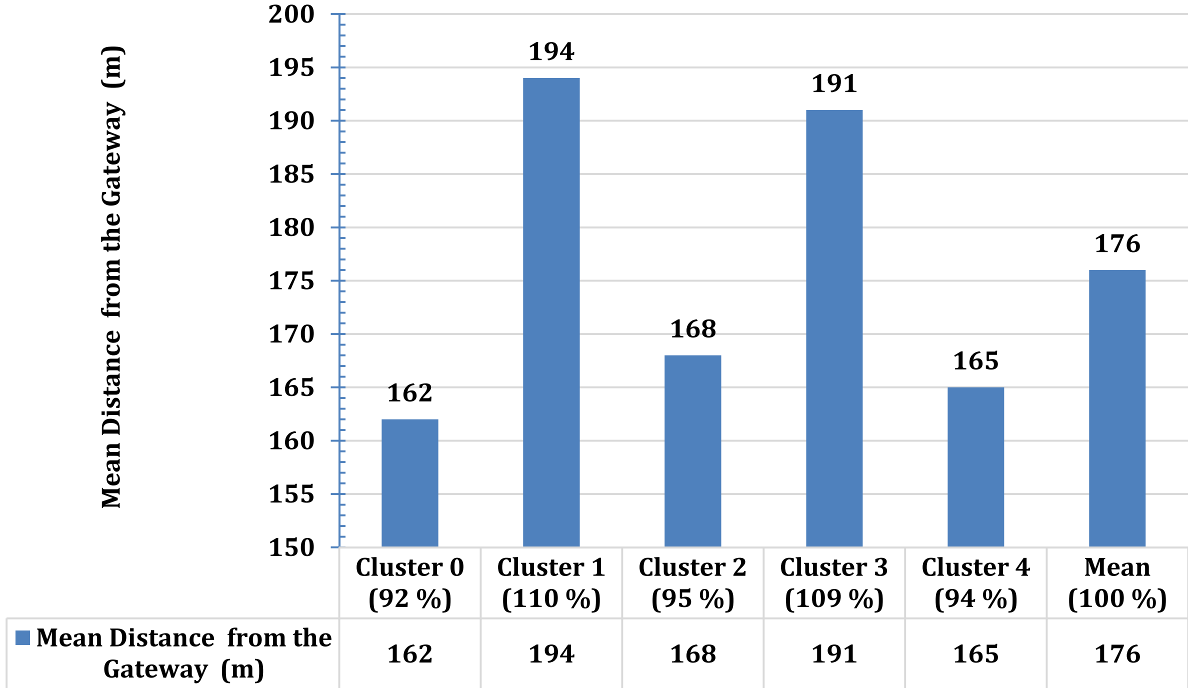


Figure 8: The bar chart of the mean distance of the sensor nodes from the gateway as determined using the enhanced fuzzy C-means

4. Conclusion

The optimum cluster and optimal gateway placement for IoT sensor network is studied. The Elbow method is employed to determine the optimum number of clusters while the enhanced fuzzy C-means is used for the optimal l=gateway placement. The results for the case study sensor network showed that 5 clusters are required based on Elbow method and mean sensor node to gateway distance of 176 m is achieved for the network.

References

1. Abner, M., Wong, P. K. Y., and Cheng, J. C. (2022). Battery lifespan enhancement strategies for edge computing-enabled wireless Bluetooth mesh sensor network for structural health

monitoring. *Automation in Construction*, *140*, 104355.

2. Bathre, M., and Das, P. K. (2023). Smart dual battery management system for expanding lifespan of wireless sensor node. *International Journal of Communication Systems*, *36*(3), e5389.
3. Bellini, P., Nesi, P., and Pantaleo, G. (2022). IoT-enabled smart cities: A review of concepts, frameworks and key technologies. *Applied Sciences*, *12*(3), 1607.
4. Bockelmann, C., Pratas N. Nikopour, H. Au, K. Svensson, T. Stefanovic, C. Popovski P. and Dekorsy A. (2016). Massive machine-type communications in 5g: physical and Mac-layer solutions. *IEEE Communications Magazine,* 54(9): 59 – 65.
5. Chataut, R., Phoummalayvane, A., and Akl, R. (2023). Unleashing the power of IoT: A comprehensive review of IoT applications and future prospects in healthcare, agriculture, smart homes, smart cities, and industry 4.0. *Sensors*, *23*(16), 7194.
6. Deniz, F., Bagci, H., Korpeoglu, I., and Yazıcı, A. (2021). Energy-efficient and fault-tolerant drone-BS placement in heterogeneous wireless sensor networks. *Wireless Networks*, *27*, 825-838.
7. El Ouadi, M. R., and Hasbi, A. (2022). Efficient organization of nodes in wireless sensor networks (clustering location-based LEACH). *International Journal of Electrical and Computer Engineering*, *12*(1), 1011.
8. Hamka, M., and Ramdhoni, N. (2022, November). K-means cluster optimization for potentiality student grouping using elbow method. In *AIP Conference Proceedings* (Vol. 2578, No. 1). AIP Publishing.
9. Islam, K. Y., Ahmad, I., Habibi, D., and Waqar, A. (2022). A survey on energy efficiency in underwater wireless communications. *Journal of Network and Computer Applications*, *198*, 103295.
10. Kaur, G., Chanak, P., and Bhattacharya, M. (2021). Energy-efficient intelligent routing scheme for IoT-enabled WSNs. *IEEE Internet of Things Journal*, *8*(14), 11440-11449.
11. Pizzo, A., Sanguinetti, L., and Marzetta, T. L. (2022). Fourier plane-wave series expansion for holographic MIMO communications. *IEEE transactions on wireless communications*, *21*(9), 6890-6905.
12. Rajkumar, Y., and Santhosh Kumar, S. V. N. (2024). A comprehensive survey on communication techniques for the realization of intelligent transportation systems in IoT based smart cities. *Peer-to-Peer Networking and Applications*, *17*(3), 1263-1308.
13. Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, *2*(5), 377.
14. Saunders, S. R., and Aragón-Zavala, A. A. (2024). *Antennas and propagation for wireless communication systems*. John Wiley & Sons.
15. Shaker Reddy, P. C., Sucharitha, Y. (2022). IoT-enabled energy-efficient multipath power control for underwater sensor networks. *International Journal of Sensors Wireless Communications and Control*, *12*(6), 478-494.
16. Shanmugam, R., and Kaliaperumal, B. (2021). An energy-efficient clustering and cross-layer-based opportunistic routing protocol (CORP) for wireless sensor network. *International Journal of Communication Systems*, *34*(7), e4752.
17. Shen, Y., Liu, Y., Tian, Y., & Na, X. (2022). Parallel sensing in metaverses: Virtual-real interactive smart systems for “6S” sensing. *IEEE/CAA Journal of AutomaticaSinica*, *9*(12), 2047-2054.
18. Zhang, Q., and Zhang, K. (2022). Protecting location privacy in iot wireless sensor networks through addresses anonymity. *Security and Communication Networks*, *2022*(1), 2440313.