K-Means Algorithm For Determination Of Optimal Gateway Placement In Internet Of Things Sensor Network

Chibuzor Henry Amadi[1]
Department of Electrical and Electronic Engineering
Imo State University Owerri, Imo State

Akaniyene Benard Obot[2]
Department of Electrical /Electronic Engineering,
University of Uyo, Akwa Ibom State, Nigeria
akaniyeneobot@uniuyo.edu.ng

Kufre Monday Udofia[3]
Department of Electrical /Electronic Engineering,
University of Uyo, Akwa Ibom State, Nigeria
kmudofia@uniuyo.edu.ng

Kingsley M. Udofia[4]
Department of Electrical /Electronic Engineering,
University of Uyo, Akwa Ibom State, Nigeria

Abstract— **In this paper, K-Means algorithm for determination of optimal gateway placement in Internet of Things Sensor network is presented. In the K-Means method, K indicates the number of clusters in the network while the K-Means algorithm is used to determine the optimal locations of each of the K gateways in the network. The detailed K-Means algorithm is presented along with sample numerical computation of the optimal location of gateways in a network with eight sensor nodes. Furthermore, more robust simulations were conducted on a 5000 nodes IoT sensor network with a square coverage area of 800 m by 800 m. Particularly, two different values of k were considered, namely K = 5 and K = 4. For K =5, the results of the coordinates of the cluster heads as determined by the K-Means algorithm were (400.54, 460.22), (180.14, 205.22), (140.11,605.32), (620.41,608.33) and (650.07,200.02) while the coordinates of the cluster heads for K = 4 are (217.09,205.19), (217.30,605.48), (608.22,608.14), and (608.10,200.22). in all, the K-Means can be effectively used to determine the optimal location of the gateways in an Internet of Things sensor network when the number of clusters in the network are known.**

Keywords— K-Means Algorithm, Optimal Gateway Placement, Internet Of Things, Sensor Network, Elbow method, Energy Consumption

1.0 Introduction

Clustering in internet of Thing (IoT) sensor networks is increasingly being used to optimize the network lifespan by minimizing the overall energy consumed by the sensors in a network [1,2,3]. Specifically, when the number of clusters is known, then a method is needed to determine the optimal location for the placement of the cluster heads or gateways [4,5,6]. The optimal location of the gateways ensures that the average energy use for the communication within the network is minimal [7,8].

Today, among several methods, K-Means is one of the methods which can be used to determine the optimal gateway location such as in a clustered IoT sensor network [9,10.11]. Apart from identifying or formulating clusters, K-Means can optimally select location for the cluster head for optimum performance [12,13]. In the K-Means algorithm, the *k* parameter which denotes the number of clusters must be specified [14,15]. In addition, computing the distances between the proposed cluster centers and the sensor nodes is an essential task to be performed in K-Means algorithm to effectively map optimal locations for the cluster heads [16,17]. The initial cluster location is randomly selected. Accordingly, in this work, the details of K-Means application in optimal gateway placement in a clustered IoT sensor network are presented along with some numerical computations and simulation results.

2. Methodology

The focus in this paper is to use K-Means to determine the optimal location of the gateway in an Internet of Things (IoT) sensor network. The procedure used to accomplish this is presented in this section.

2.1 Determination of Optimal Gateway Location using K-Means Algorithm

Consider the illustration presented in Figure 1 where raw dataset is given. First the number of clusters k must be defined, and randomly placed on the two dimensional space. Then K-Means algorithm is applied to iteratively update each of the cluster centroid. The clusters are formed based on similarities of the members around it. K-Means method of finding optimal location for cluster head is executed based on Algorithm 1:

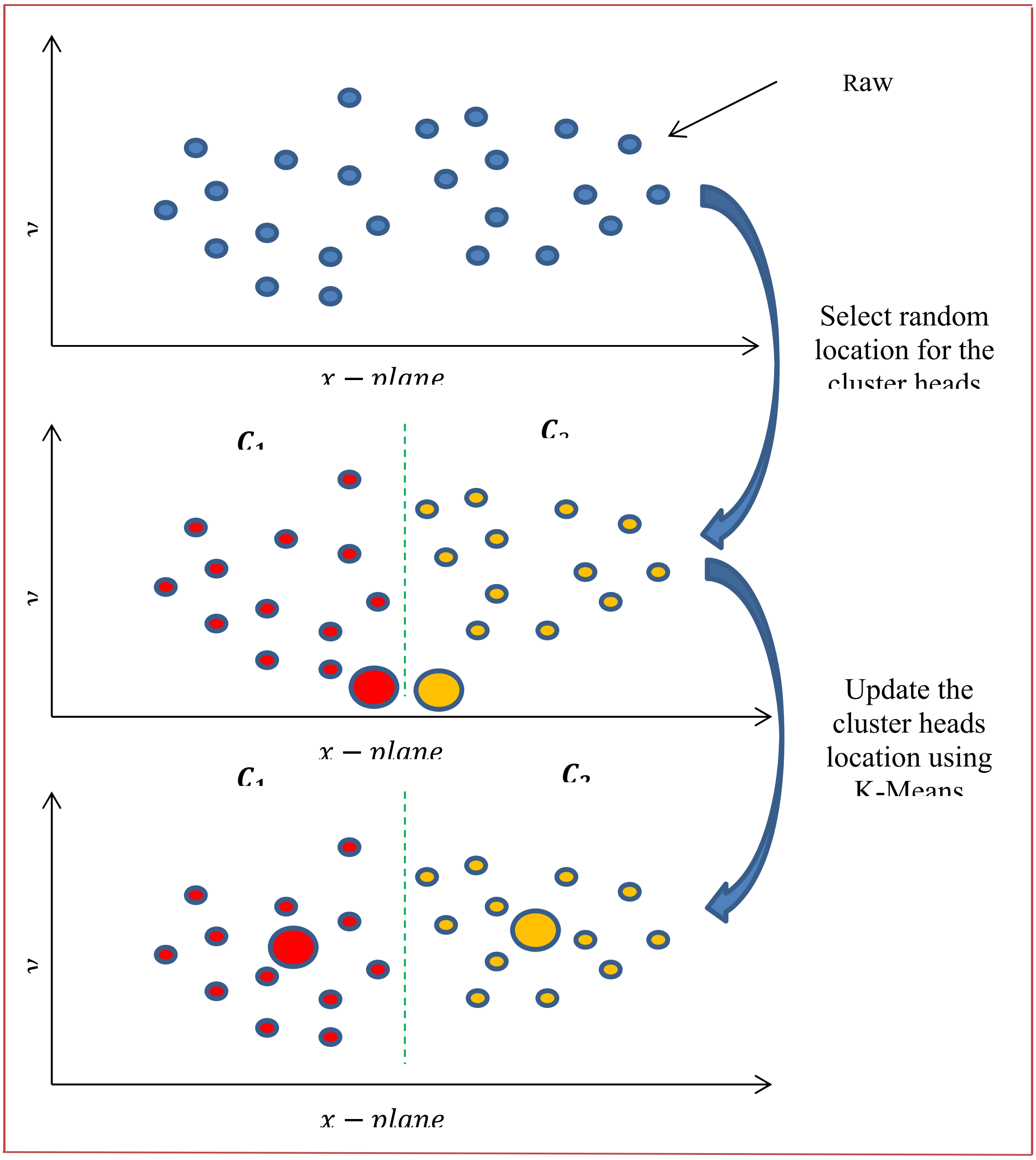

Figure 1: Cluster head selection concept for K-Means clustering method [14]

Algorithm 1: Gateway optimal location using K-Means Algorithm

1: **Begin**

2: Initialize k

3: Randomly select k points as the cluster centers

4: Compute the Euclidean distance between each data point and all the selected k

5: Assign each data point to the cluster closest to it based on the calculated distance

6: Update the new cluster center by computing the mean of all the data points within the cluster

7: **if** the center changes **then**

8: **go to** 4

9: **else**

10: **end**

2.2 Sample Numerical Computation of Optimal Gateway Locations using K-Means Algorithm

To illustrate how the cluster centroid can be determined based on the K-Means algorithm presented as Algorithm 1, we consider eight sensors located on the (x, y) plane as $S_1(2,5)$, $S_2(6,3)$, $S_3(7,1)$, $S_4(3,2)$, $S_5(4,6)$, $S_6(2,7)$, $S_7(5,7)$, and $S_8(5,5)$.

Step 1: Begin

Step 2: Let the number of clusters k be set to 3.

Step 3: Let three random locations for the cluster heads be selected as $C_1 = S_1(2,5)$, $C_2 = S_4(3,2)$, and $C_3 = S_8(5,5)$

Step 4: Compute the Euclidean distances between each of the data point and the cluster centers using $d_i = \sqrt{(C_{i1} - S_{i1})^2 + (C_{i2} - S_{i2})^2}$. This will result in the matrix shown in Table 1. The scatter plot of the sensor nodes based on Table 1 is shown in Figure 2 where the three randomly selected cluster heads are shown in red colour.

Table 1: The Euclidean distances between each of the data point and the cluster centers after the first iteration

	C_1	C_2	C_3	$Cluster$
S_1	0	3.2	3	C_1
S_2	4.5	3.2	2.2	C_3
S_3	6.4	4.1	4.5	C_2
S_4	3.2	0	3.6	C_2
S_5	2.2	4.1	1.4	C_3
S_6	2	5.1	3.6	C_1
S_7	3.6	5.4	2	C_3
S_8	3	3.6	0	C_3

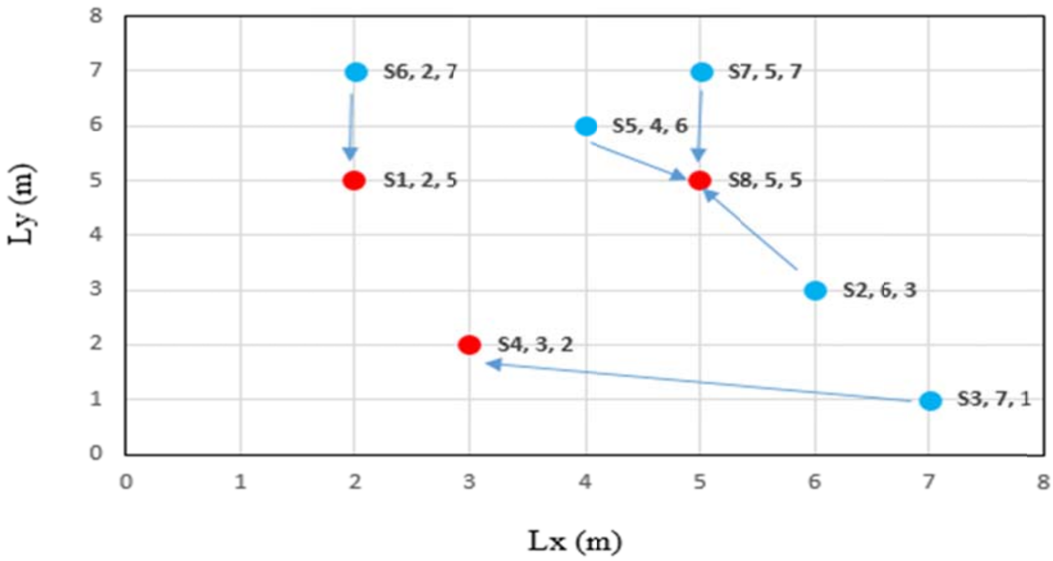

Figure 2: The scatter plot of the eight sensor nodes based on Table 1 where the three randomly selected cluster heads are shown in red colour.

Step 6: Compute the mean of all the members per cluster:

$$C_1 = \frac{S_1(2,5) + S_6(2,7)}{2} = \frac{(4,12)}{2} = (2,6)$$

$$C_2 = \frac{S_3(7,1) + S_4(3,2)}{2} = \frac{(10,3)}{2} = (5,1.5)$$

$$C_3 = \frac{S_2(6,3) + S_5(4,6) + S_7(5,7) + S_8(5,5)}{4}$$

$$= \frac{(20,21)}{4} = (5,5.3)$$

Step 7: Repeat step 4 with the new cluster centers. The matrix is presented in Table 2. Note that after the second iteration, the cluster members did not relocate; hence the iteration stops. The scatter plot of the sensor nodes based on Table 2 is shown in Figure 3 where the three randomly selected cluster heads are shown in red colour.

Table 2: The Euclidean distances between each of the data point and the cluster centers after the second iteration

Sensors	C_1	C_2	C_3	$Cluster$
S_1	1	3.2	3	C_1
S_2	5	3.2	2.2	C_3
S_3	7.1	4.1	4.5	C_2
S_4	4.1	0	3.6	C_2
S_5	2	4.1	1.4	C_3
S_6	1	5.1	3.6	C_1
S_7	3.2	5.4	2	C_3
S_8	3.2	3.6	0	C_3

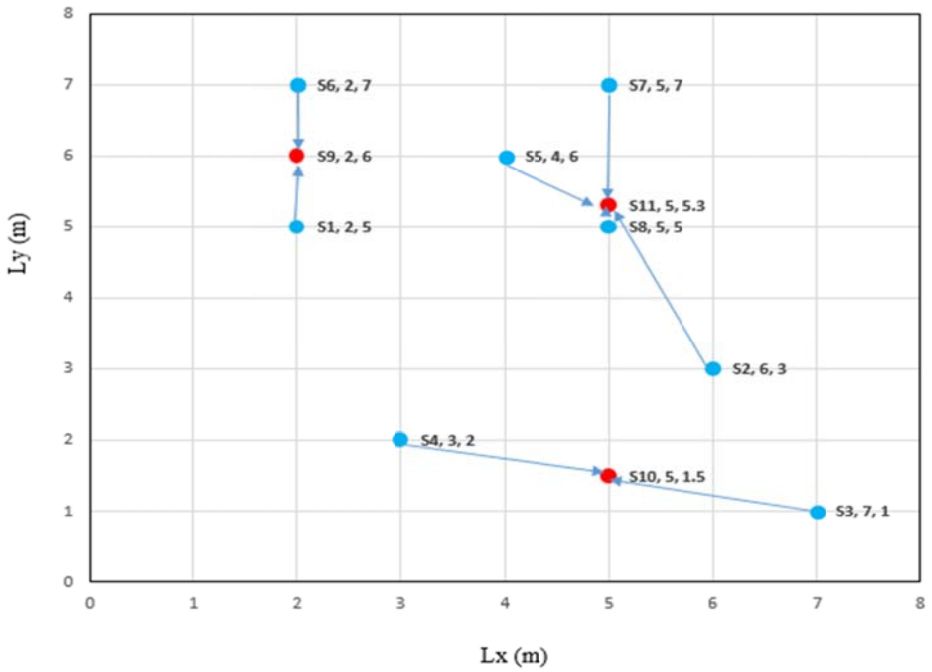

Figure 3: The scatter plot of the eight sensor nodes based on Table 2 where the three randomly selected cluster heads are shown in red colour.

3 Results and Discussion

The case study sensor network with 800 m x 800 m area and 5000 sensor nodes was used in the simulation that was based on a Python program developed in Pycharm development environment. Notably, two sets of experiments were simulated, the first case with the number of clusters (K) set to 5 and in the second case with the number of clusters (K) set to 4.

3.1 The Results of the Optimal Gateway Location using K-Means Method with K =5

In this simulation, the number of clusters (K) is set to 5 while K-Means is used to determine the cluster head location. The result of the simulation is presented in Figure 4. From the results, the coordinates of the cluster heads as determined by the K-Means algorithm are (400.54, 460.22), (180.14, 205.22), (140.11, 605.32), (620.41, 608.33) and (650.07, 200.02).

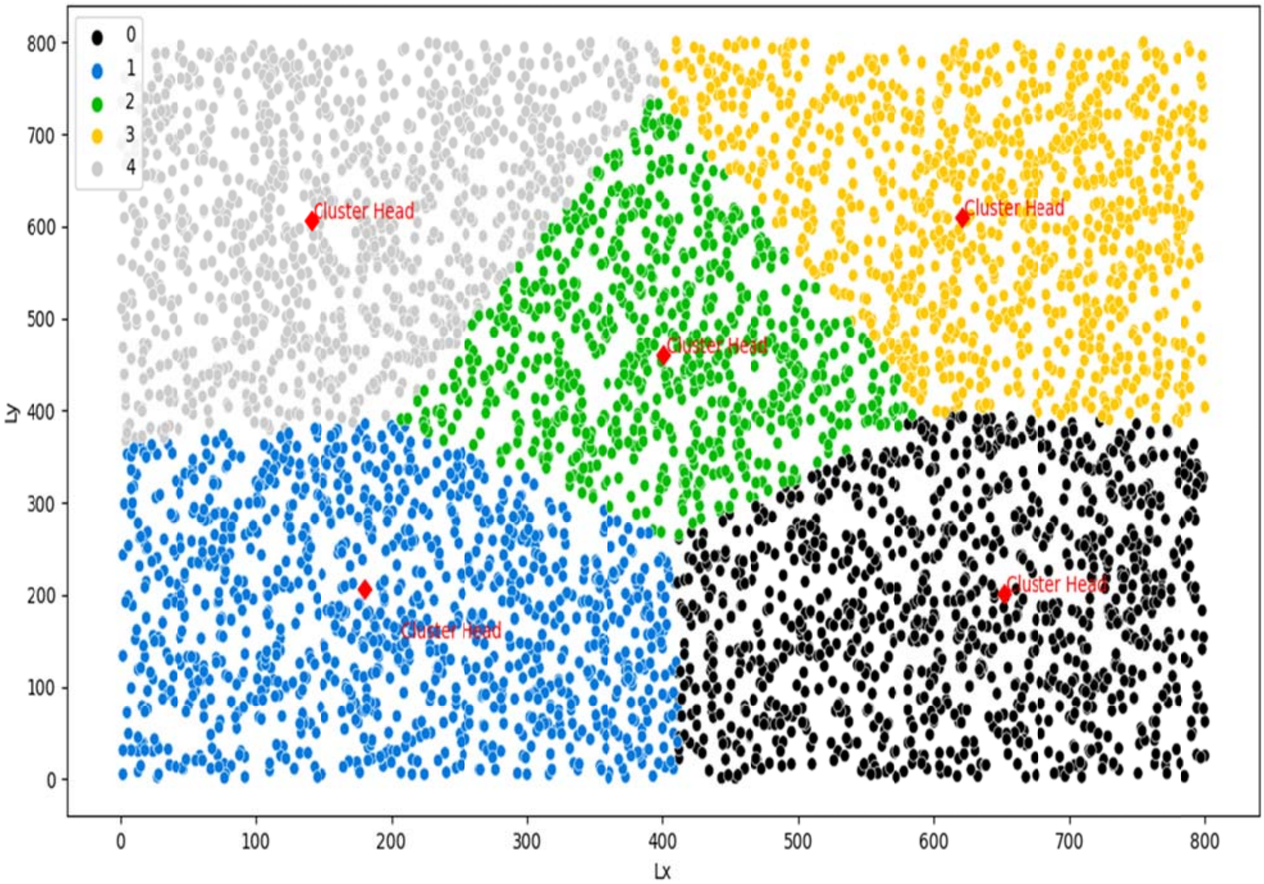

Figure 4: Optimal gateway location using K-Means with K =5

3.2 The Results of the Optimal Gateway Location using K-Means Method with K =4

In this simulation, the number of clusters (K) is set to 4 while K-Means is used to determine the cluster head location. The result of the simulation is presented in Figure 5. From the results, the coordinates of the cluster heads as determined by the K-Means algorithm are (217.09,205.19), (217.30,605.48), (608.22,608.14), and (608.10,200.22).

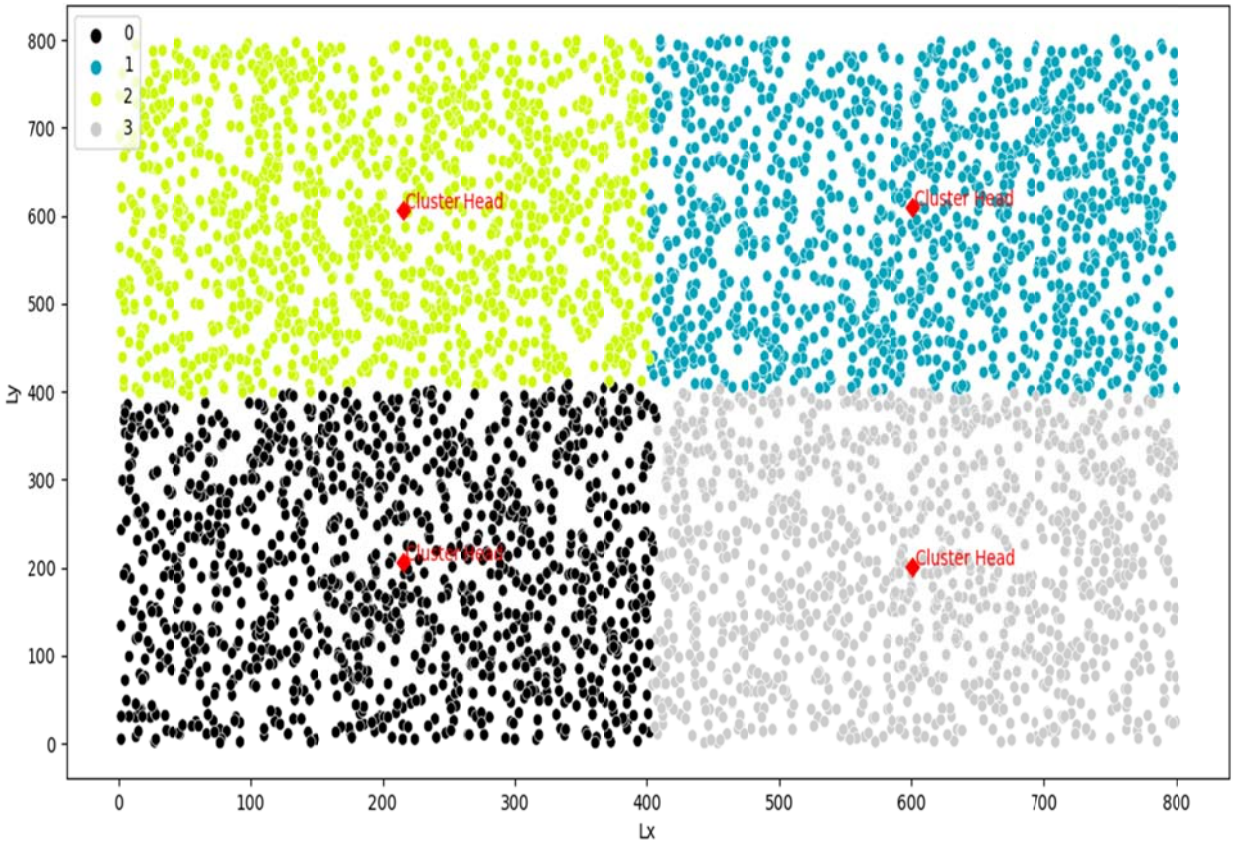

Figure 5: Optimal gateway location using K-Means with K = 4

4 Conclusion

The optimal gateway placement problem is addressed in this paper. The solution presented is based on the use of K-Means algorithm to determine the optimal location of the gateways based on the specified number of clusters. Specifically, two simulation scenario were presented for the case study sensor network; the first case with 5 clusters and the second case with 4 clusters. In each case, the K-Means algorithm is used to determine the location of the gateways in each of the clusters.

References

1. Prasad, V. K., & Periyasamy, S. (2023). energy optimization-based clustering protocols in wireless sensor networks and Internet of Things-survey. *International Journal of Distributed Sensor Networks*, *2023*.
2. Jubair, A. M., Hassan, R., Aman, A. H. M., Sallehudin, H., Al-Mekhlafi, Z. G., Mohammed, B. A., & Alsaffar, M. S. (2021). Optimization of clustering in wireless sensor networks: techniques and protocols. *Applied Sciences*, *11*(23), 11448.
3. Jothikumar, C., Ramana, K., Chakravarthy, V. D., Singh, S., & Ra, I. H. (2021). An efficient routing approach to maximize the lifetime of IoT-based wireless sensor networks in 5G and beyond. *Mobile Information Systems*, *2021*, 1-11.
4. Raghavendra, Y. M., & Mahadevaswamy, U. B. (2021). Energy efficient intra cluster gateway optimal placement in wireless sensor network. *Wireless Personal Communications*, *119*, 1009-1028.
5. Patil, S., & Gokhale, P. (2021). Throughput optimization-based gateways placement methods in wireless networks. *International Journal of Computer Networks and Applications*, *8*(4), 346-357.
6. Patil, S. (2021). Distance aware gateway placement optimization for machine-to-machine (M2M) communication in iot network. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(2), 1995-2005.
7. Raghavendra, Y. M., & Mahadevaswamy, U. B. (2021). Energy efficient intra cluster gateway optimal placement in wireless sensor network. *Wireless Personal Communications*, *119*, 1009-1028.

8. Zhou, D., Sheng, M., Wu, J., Li, J., & Han, Z. (2021). Gateway placement in integrated satellite–terrestrial networks: Supporting communications and Internet of Remote Things. *IEEE Internet of Things Journal*, *9*(6), 4421-4434.
9. Zhu, B., Bedeer, E., Nguyen, H. H., Barton, R., & Henry, J. (2020). Improved soft-K-Means clustering algorithm for balancing energy consumption in wireless sensor networks. *IEEE Internet of Things Journal*, *8*(6), 4868-4881.
10. Kadim, A. M., Al-Mukhtar, F. S., Hasan, N. A., Alnajjar, A. B., & Altaei, M. S. M. (2022). K-Means clustering of optimized wireless network sensor using genetic algorithm. *Periodicals of Engineering and Natural Sciences*, *10*(3), 276-285.
11. Li, L., Qiu, Y., & Xu, J. (2022, April). A K-Means clustered routing algorithm with location and energy awareness for underwater wireless sensor networks. In *Photonics* (Vol. 9, No. 5, p. 282). MDPI.
12. Shyjith, M. B., Maheswaran, C. P., & Reshma, V. K. (2021). Optimized and dynamic selection of cluster head using energy efficient routing protocol in WSN. *Wireless Personal Communications*, *116*, 577-599.
13. Khalifeh, A., Abid, H., & Darabkh, K. A. (2020). Optimal cluster head positioning algorithm for wireless sensor networks. *Sensors*, *20*(13), 3719.
14. Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-Means clustering algorithm. *IEEE access*, *8*, 80716-80727.
15. Chowdhury, K., Chaudhuri, D., & Pal, A. K. (2021). An entropy-based initialization method of K-Means clustering on the optimal number of clusters. *Neural Computing and Applications*, *33*, 6965-6982.
16. Zhu, B., Bedeer, E., Nguyen, H. H., Barton, R., & Henry, J. (2020). Improved soft-K-Means clustering algorithm for balancing energy consumption in wireless sensor networks. *IEEE Internet of Things Journal*, *8*(6), 4868-4881.
17. Raj, B., Ahmedy, I., Idris, M. Y. I., & Md. Noor, R. (2022). A survey on cluster head selection and cluster formation methods in wireless sensor networks. *Wireless Communications and Mobile Computing*, *2022*, 1-53.