K-MEANS-BASED ENERGY-AWARE CLUSTER HEAD SELECTION IN WIRELESS SENSOR NETWORKS

Isaac A. Ezenugu[1]
Department of Electrical/Electronic Engineering,
Imo State University (IMSU), Owerri, Nigeria
Corresponding Author:
isaac.ezenugu@yahoo.com

Dike Happiness Ugochi[2]
Department of Electrical/Electronic Engineering,
Imo State University (IMSU), Owerri, Nigeria
happiness@siat.ac.cn

***Abstract* — K-means-based energy-aware cluster head selection in wireless sensor networks (WSN) was presented in this paper. The algorithm seek to utilize WSN devices with minimal energy dissipation rate as cluster heads so that they can support more cluster slave devices in the network. The K-means algorithm in this paper used threshold energy dissipation rate of 1.4 joules for the selection of the cluster heads in the WSN with 50 WSN devices distributed randomly over a region of about 3 km by 3 km. The k-means algorithm was simulated in Matlab software. The simulation results shows that there are 4 nodes with energy dissipation rate less than 1.4Joules; as such the 4 nodes are selected as cluster heads (relays). The K-means algorithm was then used to cluster the rest of the 46 WSN nodes around the 4 cluster heads.**

Keywords— Clustering, K-Means Algorithm, Cluster Heads, Wireless Sensor Network, Wireless Sensor Devices, Sensor Node

I. INTRODUCTION

Wireless sensor networks (WSN) are networks that use autonomous sensor devices which are randomly distributed in space to monitor certain parameters in the region and transmit the data to the server [1,2,3,4,5,6]. The data is often sent to the server or storage system via a base station. Since the sensors are in most cases power-limited, efforts are always made by WSN designers at minimize the overall average energy needed to transmit the data from the sensor devices. Among other things, designers consider the expected propagation loss that the signal will suffer as it propagates within the network coverage area [7,8,9,10,11,12]. The signal strength is usually stronger for those network nodes that are closer to the base station. Hence, those nodes closer to the base station expend smaller amount of energy in the communication via the base station. As such, the D2D technologies tend to take advantage of such idea to minimize the overall energy demand of the entire network by allowing the nodes that are farther away from the base station to transmit their signal to the nodes that are closer to the base station which in turn relay the signal to the base station [13, 14, 15, 16, 17, 18]. It has been found that, apart from efficient channel allocation mechanism [19, 20, 21, 22], the D2D technology can improve the overall system performance [23, 24, 25, 26,27, 28, 29, 30, 31] .

Accordingly, in this paper, energy aware k-means algorithm-based procedure [32,33,34,35,36,37,38] is presented for clustering sensor nodes; enables cluster head selection and facilitates device-to-device communication that will minimize the overall average energy demand for the data transmission in the WSN. Salient ideas behind the k-means algorithm are presented and sample WSN was used to demonstrate the applicability of the ideas presented in this paper.

II. K-MEANS CLUSTERING ALGORITHM

K-means is one of the most popular unsupervised learning algorithms that are used for clustering data. K-Mean algorithm was proposed by MacQueen in 1967 [39,40,41,42]. The algorithm focuses at minimizing an *objective function*, which is a squared error function [43,44,45,46,47,48,49]. The objective function in K-means algorithm is given as:

$$j = \sum_{j=1}^{k}\left(\sum_{i=1}^{x}\left\|x_i^j - c_j\right\|^2\right) \quad (1)$$

Where x_i^j is a data point , c_j is the center of the cluster and $\left\|x_i^j - c_j\right\|$ is the distance between x_i^j and c_j . The value of $\left\|x_i^j - c_j\right\|$ is a measure of the distance from the cluster center to the n data points.

K-means algorithm is as follows:

1. First, K points are placed in the space occupied by the object (nodes or devices) and the points are used to represent the objects that are being clustered. The points at this instance are the initial group centroids for the clustering.
2. Each object is assigned to a group to which the object has the closest centroid
3. After assigning centroid to all the objects, the positions of the K centroids are recalculated.
4. Steps 2 and step 3 are repeated until the centroids remain the same. This produces appropriate separation distance of the objects from which the metric to be minimized is calculated.

III. SIMULATION AND RESULTS

In this paper, a wireless sensor network (WSN) that covers a region of about 3 km by 3 km with 50 WSN devices was considered. Notably, in the study, the WSN devices with the least energy dissipation rate was selected as relay nodes (cluster heads). Specifically, the node with energy dissipation rate less than 1.4Joules/ hour of transmission is considered a cluster head. The clustering of the cluster slaves to the relays was performed with k-means and the algorithm was simulated in Matlab software. Table 1 shows the position of the 50 WSN nodes in the region with their energy dissipation rate. The coordinate positions of the sensor nodes and the position of the base station is shown in Figure 1. The rate of energy dissipation plot per hour of transmission is shown in Figure 2.

The simulation results show that there are 4 nodes with energy dissipation rate less than 1.4 Joules; as such the 4 nodes are selected as cluster heads (relays). The x-y coordinate of the relays are shown in Table 2. The location of the cluster heads (relay) are presented in Figure 3. The clustering of the nodes to the relay is shown in Figure 4. Essentially, in the WSN considered, 50 WSN nodes were studied and the K-means algorithm selected 4 of the WSN nodes as the cluster heads and then the rest of the 46 WSN nodes were clustered around the 4 cluster heads. Each of the 4 WSN nodes selected as cluster heads has energy dissipation rate less than 1.4Joules.

Table 1: Position of the 50 WSN nodes in the region with their energy dissipation rate

Device number	x-coordinate (m)	y-coordinate (m)	Energy dissipation per hour of transmission
1	1043.1	1602.2	5.9706
2	449.99	269.85	6.669
3	1758.3	335.12	1.2879
4	786.44	408.88	6.5324
5	133.36	2036	4.2617
6	2264.8	1485.5	1.4458
7	728.36	569.13	5.4061
8	1327.2	1485	2.7326
9	2063.4	442.82	2.1078
10	1077.7	164.92	2.8494
11	2209	2552.1	2.3186
12	1184.1	1681.7	2.7016
13	2050.2	2788.8	1.3839
14	2112.1	2090	6.7168
15	1326.9	1748.4	3.5368
16	58.733	2446.2	5.8474
17	992.57	2637	7.2565
18	1272.9	2966.7	5.492
19	810.81	1.5671	5.8222
20	591.16	2596.3	5.0066
21	2465.2	1837.7	2.1154
22	1289.8	2969.9	5.4132
23	2663.3	1583	8.677
24	1173.5	1438.6	8.8653
25	2307.3	2404	3.4326
26	1190.4	683.53	2.8762
27	2425.5	1494.3	6.0848
28	2265.2	2702.6	6.7628
29	1132.2	1724	4.7533
30	648.06	2535.5	2.8538

31	2371.2	2215.9	9.5314
32	2847.9	1758	1.7386
33	982.7	740.2	1.9514
34	2013.8	1999.2	2.2784
35	1315.9	250.45	2.4981
36	2500.5	1877.9	6.5886
37	2306.6	1982.8	6.1634
38	501.76	2189.3	1.4687
39	2585.9	2672.3	9.3808
40	2969.6	2946.9	7.558
41	1543.3	2307.1	7.6406
42	2652.8	1744.3	1.5706
43	1764.1	2784.9	8.744
44	464.26	1740.3	9.4096
45	599.59	50.949	9.8596
46	1220.9	362.58	8.7304
47	2246.1	2588.1	8.07
48	2476.8	1452.9	5.6204
49	2369.9	2534.6	2.5984
50	955.57	628.22	4.5873

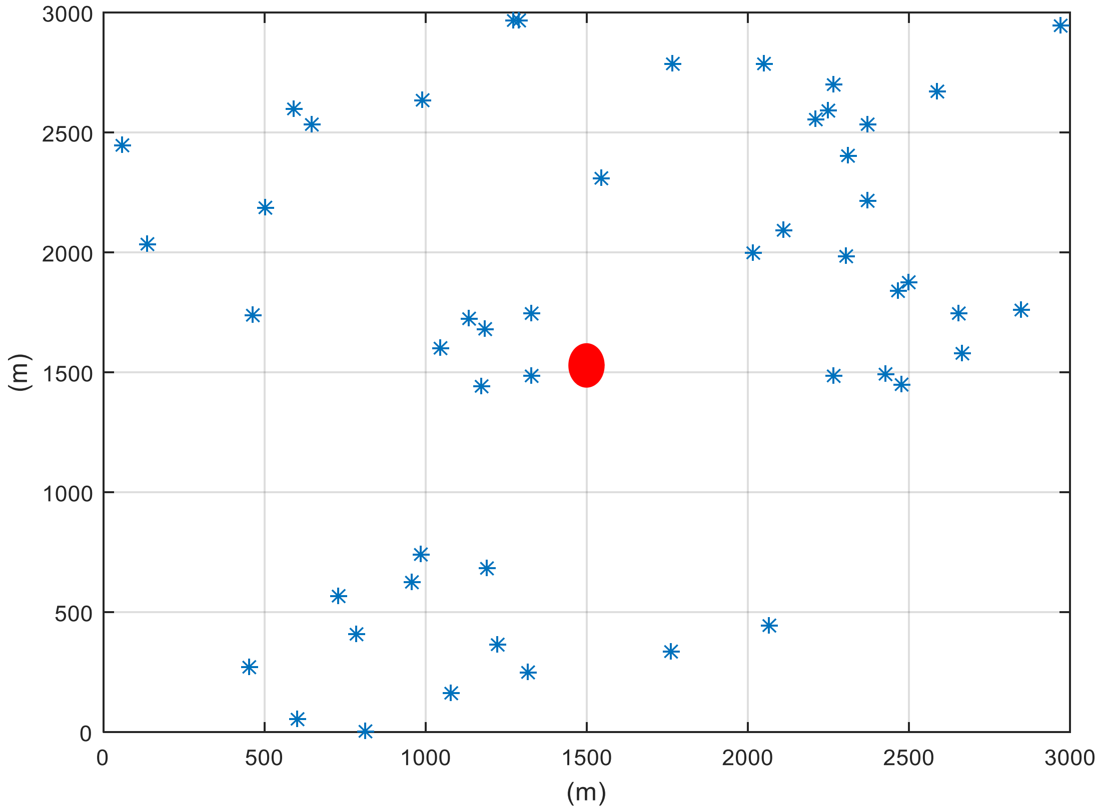

Figure 1; Coordinate positions of the WSN nodes and the cluster head heads (relay)

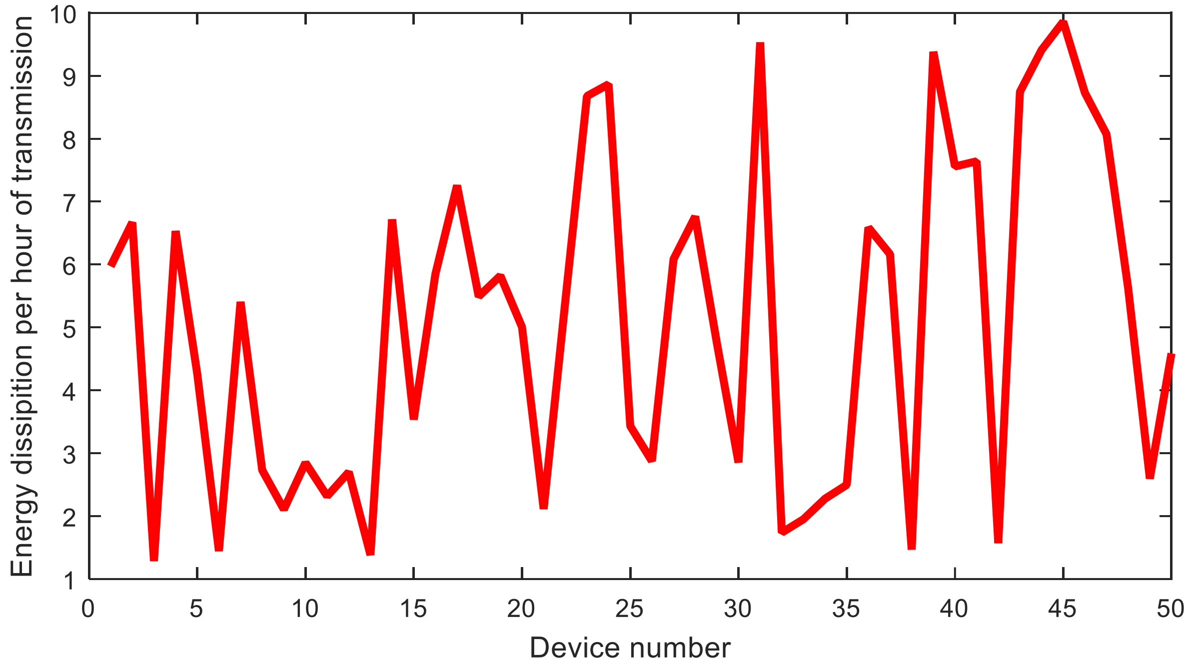

Figure 2; Energy dissipation per hour of transmission.

Table 2. The x-y coordinate of the relays and their energy dissipation per hour of transmission

Relay number (Device number)	x-coordinate (m)	y-coordinate (m)	Energy dissipation per hour of transmission
1(3)	1758.3	335.12	1.2879
2(6)	2264.8	1485.5	1.4458
3(13)	2050.2	2788.8	1.3839
4(38)	501.76	2189.3	1.4687

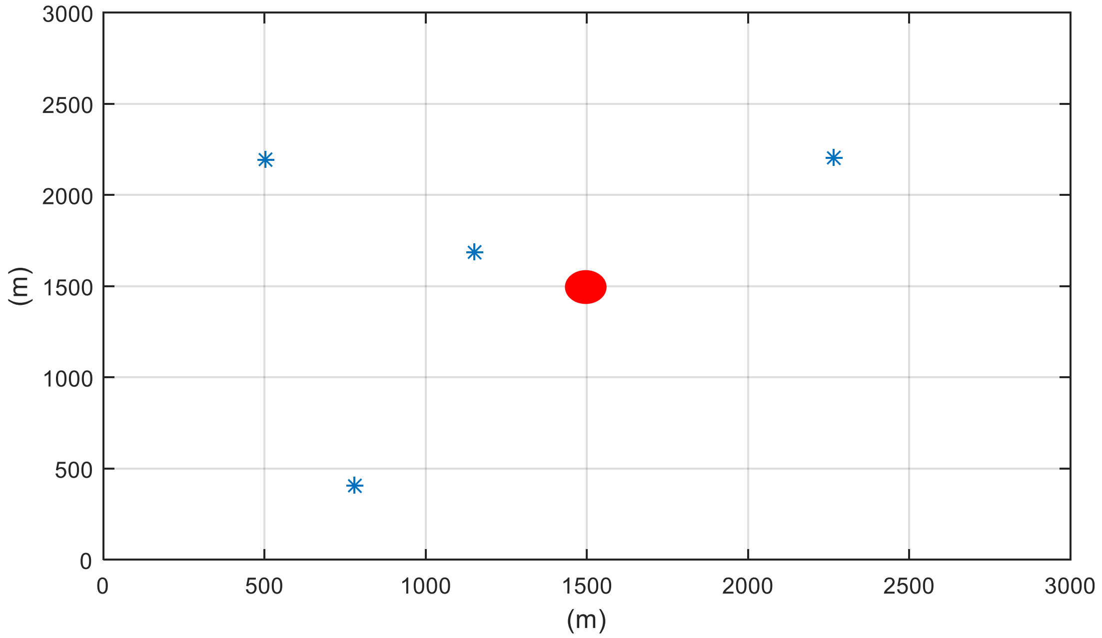

Figure 3; Position of the cluster heads (relay)

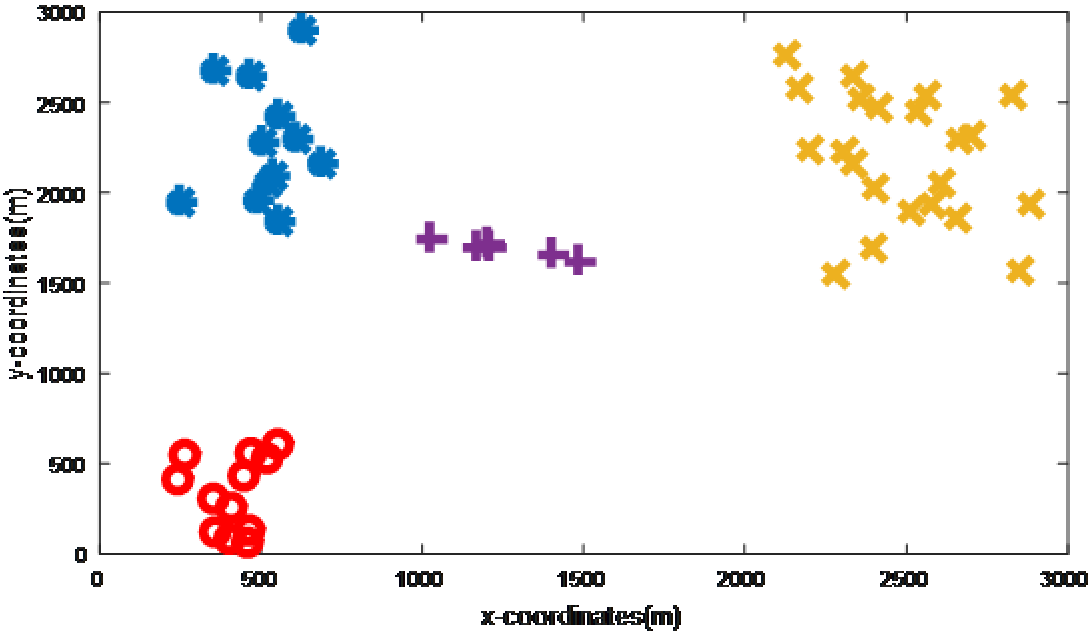

Figure 4; Clustering of the WSN nodes using k-means algorithm.

IV. CONCLUSIONS

Clustering of wireless sensor network devices was presented. The clustering was based on K-means algorithm which used threshold energy dissipation rate of the WSN devices as the key parameter for the selection of the cluster heads. The key ideas and mathematical expression about K-means algorithm were presented and a sample WSN was used to demonstrate the application of the K-means algorithm in clustering of WSN devices.

REFERENCES

1. Carlos-Mancilla, M., López-Mellado, E., & Siller, M. (2016). Wireless sensor networks formation: approaches and techniques. *Journal of Sensors*, *2016*.
2. Ali, A., Ming, Y., Chakraborty, S., & Iram, S. (2017). A comprehensive survey on real-time applications of WSN. *Future internet*, *9*(4), 77.
3. Xu, G., Shen, W., & Wang, X. (2014). Applications of wireless sensor networks in marine environment monitoring: A survey. *Sensors*, *14*(9), 16932-16954.
4. Pule, M., Yahya, A., & Chuma, J. (2017). Wireless sensor networks: A survey on monitoring water quality. *Journal of applied research and technology*, *15*(6), 562-570.
5. Dinh, N. T., & Kim, Y. (2019). Auto-Configuration in Wireless Sensor Networks: A Review. *Sensors*, *19*(19), 4281.
6. Bokare, M., & Ralegaonkar, A. (2012). Wireless sensor network: A promising approach for distributed sensing tasks. *Excel Journal of Engineering Technology and Management Science*, *1*(1), 1-9.
7. Ozuomba, S., Enyenihi, J., & Rosemary, N. C. (2018). Characterisation of Propagation Loss for a 3G Cellular Network in a Crowded Market Area Using CCIR Model. *Review of Computer Engineering Research*, *5*(2), 49-56.
8. Ozuomba, Simeon (2019) ANALYSIS OF EFFECTIVE TRANSMISSION RANGE BASED ON HATA MODEL FOR WIRELESS SENSOR NETWORKS IN THE C-BAND AND KU-BAND, Journal of Multidisciplinary Engineering Science and Technology (JMEST)
9. Akaninyene B. Obot , Ozuomba Simeon and Kingsley M. Udofia (2011); *"Determination Of Mobile Radio Link Parameters Using The Path Loss Models"* NSE Technical Transactions , A Technical Journal of The Nigerian Society Of Engineers, Vol. 46, No. 2 , April - June 2011 , PP 56 – 66.
10. Njoku Chukwudi Aloziem, Ozuomba Simeon, Afolayan J. Jimoh (2017) Tuning and Cross Validation of Blomquist-Ladell Model for Pathloss Prediction in the GSM 900 Mhz Frequency Band , *International Journal of Theoretical and Applied Mathematics*, *Journal*: Environmental and Energy Economics
11. Ozuomba, Simeon (2019) EVALUATION OF OPTIMAL TRANSMISSION RANGE OF WIRELESS SIGNAL ON DIFFERENT TERRAINS BASED ON ERICSSON PATH LOSS MODEL
12. Akaninyene B. Obot , Ozuomba Simeon and Afolanya J. Jimoh (2011); "*Comparative Analysis Of Pathloss Prediction Models For Urban Macrocellular" Nigerian Journal of Technology (NIJOTECH)* Vol. 30, No. 3 , October 2011 , PP 50 – 59
13. Rhim, H., Tamine, K., Abassi, R., Sauveron, D., & Guemara, S. (2018). A multi-hop graph-based approach for an energy-efficient routing protocol in wireless sensor networks. *Human-centric Computing and Information Sciences*, *8*(1), 1-21.

14. Elshrkawey, M., Elsherif, S. M., & Wahed, M. E. (2018). An enhancement approach for reducing the energy consumption in wireless sensor networks. *Journal of King Saud University-Computer and Information Sciences*, *30*(2), 259-267.
15. Somauroo, A., & Bassoo, V. (2019). Energy-efficient genetic algorithm variants of PEGASIS for 3D wireless sensor networks. *Applied Computing and Informatics*.
16. Khan, R. A., Mohammadani, K. H., Soomro, A. A., Hussain, J., Khan, S., Arain, T. H., & Zafar, H. (2018). An energy efficient routing protocol for wireless body area sensor networks. *Wireless Personal Communications*, *99*(4), 1443-1454.
17. Thangaramya, K., Kulothungan, K., Logambigai, R., Selvi, M., Ganapathy, S., & Kannan, A. (2019). Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT. *Computer Networks*, *151*, 211-223.
18. Behera, T. M., Mohapatra, S. K., Samal, U. C., & Khan, M. S. (2019). Hybrid heterogeneous routing scheme for improved network performance in WSNs for animal tracking. *Internet of Things*, *6*, 100047.
19. Akpan, N. E., Akpan, U. I., & Umoh, J. B. Mitigating multimedia call failure in CDMA networks.
20. Itoro J. Akpan, Ozuomba Simeon and Kalu Constance. “DEVELOPMENT OF A GUARD CHANNEL-BASED PRIORITIZED HANDOFF SCHEME WITH CHANNEL BORROWING MECHANISM FOR CELLULAR NETWORKS”. Journal of Multidisciplinary Engineering Science and Technology (JMEST)
21. Akpan, U., Kalu, C., Ozuomba, S., & Obot, A. (2013). Development of Improved Scheme for Minimising Handoff Failure due to Poor Signal Quality. *International Journal of Engineering Research and Technology*, *2*(10), 2764-2771.
22. Omini Samuel , Kalu Constance and Ozuomba Simeon , “DEVELOPMENT OF FUZZY INFERENCE SYSTEM-BASED ANALYSIS OF NON-PRIORITIZED HANDOFF SCHEME WITH STATISTICALLY MODELLED IDLECHANNEL BORROWINGMECHANISM FOR CELLULAR NETWORKS” International Multilingual Journal of Science and Technology (IMJST)
23. Gandotra, P., & Jha, R. K. (2016). Device-to-device communication in cellular networks: A survey. *Journal of Network and Computer Applications*, *71*, 99-117.
24. Kar, U. N., & Sanyal, D. K. (2018). An overview of device-to-device communication in cellular networks. *ICT express*, *4*(4), 203-208.
25. Della Penda, D. (2018). *Device-to-Device Communication in Future Cellular Networks: Resource allocation and mode selection* (Doctoral dissertation, Kungliga Tekniska högskolan).
26. Huang, J., & Gharavi, H. (2018). Performance analysis of relay-based two-way D2D communications with network coding. *IEEE transactions on vehicular technology*, *67*(7), 6642-6646.
27. Alemaishat, S., Saraereh, O. A., Khan, I., & Choi, B. J. (2019). An efficient resource allocation algorithm for D2D communications based on NOMA. *IEEE Access*, *7*, 120238-120247.
28. Safdar, G. A., Ur-Rehman, M., Muhammad, M., Imran, M. A., & Tafazolli, R. (2016). Interference mitigation in D2D communication underlaying LTE-A network. *IEEE Access*, *4*, 7967-7987.
29. Ibrahim, R. (2019). *Use of Device-to-Device communications for efficient cellular networks* (Doctoral dissertation).
30. Saraereh, O. A., Mohammed, S. L., Khan, I., Rabie, K., & Affess, S. (2019). An Efficient Resource Allocation Algorithm for Device-To-Device Communications. *Applied Sciences*, *9*(18), 3816.
31. Rodziewicz, M. (2017). Outage probability of device-to-device communications in frequency reuse-1 networks. *Mobile Networks and Applications*, *22*(6), 1058-1064.
32. Kumar, A., Jois, A. S., & Ramesh, T. K. (2019, April). Distance and Energy Aware Device to Device Communication. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0522-0525). IEEE.
33. Nwaogu Onyeyirichi E., Ozuomba Simeon and Kalu Constance , “Development of Self-Organizing Map Clustering Algorithm for Relay Selection in High Density Long Term Evolution Networks”, Science and Technology Publishing (SCI & TECH)
34. Yin, L., Gui, J., & Zeng, Z. (2019). Improving energy efficiency of multimedia content dissemination by adaptive clustering and D2D multicast. *Mobile Information Systems*, *2019*.
35. Florence Kingsley Atakpo , Ozuomba Simeon and Stephen Bliss Utibe-Abasi, “A COMPARATIVE ANALYSIS OF SELF-ORGANIZING MAP AND K-MEANS MODELS FOR SELECTION OF CLUSTER HEADS IN OUT-OF-BAND DEVICE-TO-DEVICE COMMUNICATION “Journal of Multidisciplinary Engineering Science Studies (JMESS)
36. Zhang, Y., Zhang, Y., Guo, D., & Song, M. (2016, September). An energy-saving algorithm based on base station sleeping in multi-hop D2D communication. In *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)* (pp. 1-5). IEEE.
37. Ozuomba Simeon “APPLICATION OF K-MEANS CLUSTERING ALGORITHM FOR SELECTION OF RELAY NODES IN WIRELESS SENSOR NETWORK” Journal of Multidisciplinary Engineering Science and Technology (JMEST)

38. Li, G., Xu, S., Wu, J., & Ding, H. (2018). Resource scheduling based on improved spectral clustering algorithm in edge computing. *Scientific Programming*, *2018*.
39. Morissette, L., & Chartier, S. (2013). The k-means clustering technique: General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, *9*(1), 15-24.
40. Bock, H. H. (2008). Origins and extensions of the k-means algorithm in cluster analysis. *Electronic Journal for History of Probability and Statistics*, *4*(2), 1-18.
41. Yedla, M., Pathakota, S. R., & Srinivasa, T. M. (2010). Enhancing K-means clustering algorithm with improved initial center. *International Journal of computer science and information technologies*, *1*(2), 121-125.
42. Ullman, J. D., Poggio, T., Harari, D., Zysman, D., & Seibert, D. (2012). Clustering. *Mining of Massive Datasets*, 241-280.
43. Ortega, J. P., Del, M., Rojas, R. B., & Somodevilla, M. J. (2009, March). Research issues on k-means algorithm: An experimental trial using matlab. In *CEUR workshop proceedings: semantic web and new technologies* (pp. 83-96).
44. Nishida, J. K. (2015). Estimating Single and Multiple Target Locations using K-means Clustering with Radio Tomographic Imaging in Wireless Sensor Networks.
45. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, *31*(8), 651-666.
46. Rakhlin, A., & Caponnetto, A. (2007). Stability of $ k $-means clustering. In *Advances in neural information processing systems* (pp. 1121-1128).
47. Tang, W., Yang, Y., Zeng, L., & Zhan, Y. (2019). Size Constrained Clustering With MILP Formulation. *IEEE Access*, *8*, 1587-1599.
48. Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., & Zeebaree, S. R. (2017). Combination of K-means clustering with Genetic Algorithm: A review. *International Journal of Applied Engineering Research*, *12*(24), 14238-14245.
49. Fränti, P., & Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats?. *Pattern Recognition*, *93*, 95-112.